DOI:
10.1038/s41588-018-0127-7

Document Version
Peer reviewed version

Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error

Authors

Milly S. Tedja[1,2,79], Robert Wojciechowski[3-5,79], Pirro G. Hysi[6,79], Nicholas Eriksson[7,79], Nicholas A. Furlotte[7,79], Virginie J.M. Verhoeven[1,2,8,79], Adriana I. Iglesias[1,2,8], Magda A. Meester-Smoor[1,2], Stuart W. Tompson[9], Qiao Fan[10], Anthony P. Khawaja[11,12], Ching-Yu Cheng[10,13], René Höhn[14,15], Kenji Yamashiro[16], Adam Wenocur[17], Clare Grazal[17], Toomas Haller[18], Andres Metspalu[18], Juho Wedenoja[19,20], Jost B. Jonas[21,22], Ya Xing Wang[22], Jing Xie[23], Paul Mitchell[24], Paul J. Foster[12], Barbara E.K. Klein[9], Ronald Klein[9], Andrew D. Paterson[25], S. Mohsen Hosseini[25], Rupal L. Shah[26], Cathy Williams[27], Yik Ying Teo[28,29], Yih Chung Tham[13], Preeti Gupta[30], Wanting Zhao[10,31], Yuan Shi[31], Woei-Yuh Saw[32], E-Shyong Tai[29], Xue Ling Sim[29], Jennifer E. Huffman[33], Ozren Polašek[34], Caroline Hayward[33], Goran Bencic[35], Igor Rudan[36], James F. Wilson[33,36], CREAM[#], 23andMe Research Team[*], UK Biobank Eye and Vision Consortium, Peter K. Joshi[36], Akitaka Tsujikawa[16], Fumihiko Matsuda[37], Kristina N. Whisenhunt[9], Tanja Zeller[38], Peter J. van der Spek[39], Roxanna Haak[39], Hanne Meijers-Heijboer[40,41], Elisabeth M. van Leeuwen[1,2], Sudha K. Iyengar[42-44], Jonathan H. Lass[42,43], Albert Hofman[2,45,46], Fernando Rivadeneira[2,46,47], André G. Uitterlinden[2,46,47], Johannes R. Vingerling[1], Terho Lehtimäki[48,49], Olli T. Raitakari[50,51], Ginevra Biino[52], Maria Pina Concas[53], Tae-Hwi Schwantes-An[4,54], Robert P. Igo, Jr.[42], Gabriel Cuellar-Partida[55], Nicholas G. Martin[56], Jamie E. Craig[57], Puya Gharahkhani[55], Katie M. Williams[6], Abhishek Nag[58], Jugnoo S. Rahi[12,59,60], Phillippa M. Cumberland[59], Cécile Delcourt[61], Céline Bellenguez[62-64], Janina S. Ried[65], Arthur A. Bergen[40,66,67], Thomas Meitinger[68,69], Christian Gieger[65], Tien Yin Wong[70,71], Alex W. Hewitt[23,72,73], David A. Mackey[23,72,73], Claire L. Simpson[4,74], Norbert Pfeiffer[15], Olavi Pärssinen[75,76], Paul N. Baird[23], Veronique Vitart[33], Najaf Amin[2], Cornelia M. van Duijn[2], Joan E. Bailey-Wilson[4], Terri L. Young[9], Seang-Mei Saw [29,77], Dwight Stambolian[17], Stuart MacGregor[55], Jeremy A. Guggenheim[26,80], Joyce Y. Tung[7,80], Christopher J. Hammond[6,80], Caroline C.W. Klaver[1,2,78,80]

Affiliations

1. Department of Ophthalmology, Erasmus Medical Center, Rotterdam, The Netherlands.
2. Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands.
3. Department of Epidemiology and Medicine, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.

4. Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.
5. Wilmer Eye Institute, Johns Hopkins Medical Institutions, Baltimore, Maryland, USA.
6. Section of Academic Ophthalmology, School of Life Course Sciences, King's College London, London, UK.
7. 23andMe, Inc., Mountain View, California, USA.
8. Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.
9. Department of Ophthalmology and Visual Sciences, University of Wisconsin–Madison, Madison, Wisconsin, USA.
10. Centre for Quantitative Medicine, DUKE-National University of Singapore, Singapore.
11. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
12. NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK.
13. Ocular Epidemiology Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
14. Department of Ophthalmology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland.
15. Department of Ophthalmology, University Medical Center Mainz, Mainz, Germany.
16. Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan.
17. Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
18. Estonian Genome Center, University of Tartu, Tartu, Estonia.
19. Department of Ophthalmology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.
20. Department of Public Health, University of Helsinki, Helsinki, Finland.
21. Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University of Heidelberg, Mannheim, Germany.
22. Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China.
23. Centre for Eye Research Australia, Ophthalmology, Department of Surgery, University of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Australia.
24. Department of Ophthalmology, Centre for Vision Research, Westmead Institute for Medical Research, University of Sydney, Sydney, Australia.
25. Program in Genetics and Genome Biology, Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada.
26. School of Optometry & Vision Sciences, Cardiff University, Cardiff, UK.
27. Department of Population Health Sciences, Bristol Medical School, Bristol, UK.
28. Department of Statistics and Applied Probability, National University of Singapore, Singapore.
29. Saw Swee Hock School of Public Health, National University Health Systems, National University of Singapore, Singapore.
30. Department of Health Service Research, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
31. Statistics Support Platform, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
32. Life Sciences Institute, National University of Singapore, Singapore.
33. MRC Human Genetics Unit, MRC Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh, UK.
34. Faculty of Medicine, University of Split, Split, Croatia.
35. Department of Ophthalmology, Sisters of Mercy University Hospital, Zagreb, Croatia.

36. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK.
37. Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan.
38. Clinic for General and Interventional Cardiology, University Heart Center Hamburg, Hamburg, Germany.
39. Department of Bioinformatics, Erasmus Medical Center, Rotterdam, The Netherlands.
40. Department of Clinical Genetics, Academic Medical Center, Amsterdam, The Netherlands.
41. Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands.
42. Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, USA.
43. Department of Ophthalmology and Visual Sciences, Case Western Reserve University and University Hospitals Eye Institute, Cleveland, Ohio, USA.
44. Department of Genetics, Case Western Reserve University, Cleveland, Ohio, USA.
45. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA.
46. Netherlands Consortium for Healthy Ageing, Netherlands Genomics Initiative, the Hague, the Netherlands.
47. Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands.
48. Department of Clinical Chemistry, Finnish Cardiovascular Research Center-Tampere, Faculty of Medicine and Life Sciences, University of Tampere.
49. Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere, Tampere, Finland.
50. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland.
51. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland.
52. Institute of Molecular Genetics, National Research Council of Italy, Sassari, Italy.
53. Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", Trieste, Italy.
54. Department of Medical and Molecular Genetics, Indiana University , School of Medicine, Indianapolis, Indiana, USA.
55. Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
56. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
57. Department of Ophthalmology, Flinders University, Adelaide, Australia.
58. Department of Twin Research and Genetic Epidemiology, King's College London, London, UK.
59. Great Ormond Street Institute of Child Health, University College London, London, UK.
60. Ulverscroft Vision Research Group, University College London, London, UK.
61. Université de Bordeaux, Inserm, Bordeaux Population Health Research Center, team LEHA, UMR 1219, F-33000 Bordeaux, France.
62. Institut Pasteur de Lille, Lille, France.
63. Inserm, U1167, RID-AGE - Risk factors and molecular determinants of aging-related diseases, Lille, France.
64. Université de Lille, U1167 - Excellence Laboratory LabEx DISTALZ, Lille, France.
65. Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany.
66. Department of Ophthalmology, Academic Medical Center, Amsterdam, The Netherlands.
67. The Netherlands Institute for Neurosciences (NIN-KNAW), Amsterdam, The Netherlands.
68. Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany.
69. Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.
70. Academic Medicine Research Institute, Singapore.
71. Retino Center, Singapore National Eye Centre, Singapore, Singapore.

72. Department of Ophthalmology, Menzies Institute of Medical Research, University of Tasmania, Hobart, Australia.
73. Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western Australia, Perth, Australia.
74. Department of Genetics, Genomics and Informatics, University of Tennessee Health Sciences Center, Memphis, Tenessee.
75. Department of Ophthalmology, Central Hospital of Central Finland, Jyväskylä, Finland.
76. Gerontology Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland.
77. Myopia Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
78. Department of Ophthalmology, Radboud University Medical Center, Nijmegen, The Netherlands.
79. These authors contributed equally to this work.
80. These authors jointly directed this work.

[#] A full list of CREAM consortium members appears at the end of the paper.

* Members of the 23andMe Research Team: Michelle Agee, Babak Alipanahi, Adam Auton, Robert K. Bell, Katarzyna Bryc, Sarah L. Elson, Pierre Fontanillas, David A. Hinds, Jennifer C. McCreight, Karen E. Huber, Aaron Kleinman, Nadia K. Litterman, Matthew H. McIntyre, Joanna L. Mountain, Elizabeth S. Noblin, Carrie A.M. Northover, Steven J. Pitts, J. Fah Sathirapongsasuti, Olga V. Sazonova, Janie F. Shelton, Suyash Shringarpure, Chao Tian, Vladimir Vacic, Catherine H. Wilson.

Corresponding author

Prof. dr. Caroline C.W. Klaver, Erasmus Medical Center, room Na-2808, PO Box 2040, 3000 CA, Rotterdam, the Netherlands, e-mail: c.c.w.klaver@erasmusmc.nl, phone number +31651934491, fax number +31107044657

ABSTRACT

Refractive errors, including myopia, are the most frequent eye disorders worldwide and an increasingly common cause of blindness. This genome-wide association meta-analysis in 160,420 participants and replication in 95,505 participants, increased the established independent signals from 37 to 161 and revealed high genetic correlation between Europeans and Asians (>0.78). Expression experiments and comprehensive *in silico* analyses identified retinal cell physiology and light processing as prominent mechanisms, and functional contributions to refractive error development in all cell types of the neurosensory retina, retinal pigment epithelium, vascular endothelium and extracellular matrix. Newly identified genes elicited novel mechanisms such as rod and cone bipolar synaptic neurotransmission, anterior segment morphology, and angiogenesis. Thirty-one loci resided in or near regions transcribing small RNAs, suggesting a role for post-transcriptional regulation. Our results support the notion that refractive errors are caused by a light-dependent retina-to-sclera signaling cascade, and delineate potential pathobiological molecular drivers.

INTRODUCTION

Refractive errors are common optical aberrations determined by mismatches in the focusing power of the cornea, lens and axial length of the eye. Their distribution is rapidly shifting towards myopia, or nearsightedness, all over the world. The myopia boom is particularly prominent in urban East Asia where up to 95% of twenty-year-olds in cities such as Seoul and Singapore have this refractive error[1-4]. Myopia prevalence is also rising throughout Western Europe and the USA, affecting ~50% of young adults in these regions[5,6]. While refractive errors can be optically corrected, even at moderate values they carry significant risk of ocular complications with high economic burden[7-9]. One in three individuals with high myopia (-6 diopters or worse) will develop irreversible visual impairment or blindness, mostly due to myopic macular degeneration, retinal detachment, or glaucoma[10,11]. At the other extreme, high hyperopia predisposes to strabismus, amblyopia and angle-closure glaucoma[10,12].

Refractive errors result from a complex interplay of lifestyle and genetic factors. The most established lifestyle factors for myopia are high education, lack of outdoor exposure, and excessive near work[3]. Recent research has identified many genetic variants for refractive errors, myopia, and axial length[13-25]. Two large studies, the international Consortium for Refractive Error and Myopia (CREAM)[26] and the personal genomics company 23andMe, Inc.[17,27] have provided the most comprehensive results. [28]

Given that only 3.6% of the variance of the refractive error trait was explained by the identified genetic variants[26], we presumed a high missing heritability. We therefore combined CREAM and 23andMe, and expanded the study sample to 160,420 individuals from a mixed ancestry population with quantitative information on refraction for a genome-wide association (GWAS) meta-analysis. Index variants were tested for replication in an independent cohort consisting of 95,505 individuals from the UK Biobank. We conducted systematic comparisons to assess differences in genetic inheritance and distribution of risk variants between Europeans and Asians. Polygenic risk analyses were performed to evaluate the contribution of the identified variants to the risk of myopia and hyperopia. Finally, we

integrated expression data and bioinformatics on the identified genes to gain insight into the possible mechanisms underlying the genetic associations.

RESULTS

Susceptibility loci for refractive error

We performed a GWAS meta-analysis on adult untransformed spherical equivalent (SphE) using summary statistics from 37 studies from CREAM and on age of diagnosis of myopia (AODM) from two cohorts from 23andMe (Supplementary Figure 1, Supplementary Table 1a)[26,27]. Analyses were based on ~11 million genetic variants (SNPs, insertions and deletions) genotyped or imputed to 1000 Genomes Project Phase I reference panel (version 3, March 2012 release[29]) that passed extensive quality control (Supplementary Figures 2-4, Supplementary Table 1b).

Meta-analyses were conducted in three stages: *Stage 1* CREAM (CREAM-EUR, *N*=44,192; CREAM-ASN, *N*=11,935); *Stage 2* 23andMe (*N*=104,293; Online Methods); *Stage 3* joint meta-analysis of Stage 1 and 2. As CREAM and 23andMe applied different phenotype measures, we used signed Z-scores as the mean per-allele effect size and assigned equal weights to CREAM and 23andMe. We identified 7,967 genome-wide significant genetic variants clustering in 140 loci (Figure 1a,b; Supplementary Figure 5-6, Supplementary Table 2-5, Supplementary Data 1-2), replicating all 37 previously discovered loci and finding 104 novel loci. We applied genomic control at each stage and checked for population stratification using LD score regression[30] (Stage 1-2 inflation factors (λGC) <1.1 and LD score regression intercepts ($LDSC_{intercept}$) 0.892-1.023; Supplementary Table 6; Supplementary Figure 6- 7). At Stage 3, we observed a genomic inflation (λGC=1.129; Supplementary Figure 6), probably due to true polygenicity rather than population stratification or cryptic relatedness[31]. $LDSC_{intercept}$ remained undetermined due to mixed ancestry.

To detect the presence of multiple independent signals at the discovered loci, a stepwise conditional analysis was performed with GCTA-COJO[32] on summary statistics from all European cohorts (*N*=148,485*)* using the Rotterdam Study I-III (RS I-III) as a reference panel for LD structure ($N_{RSI-III}$=10,775). This analysis yielded 27 additional independent variants, resulting in a total of 167 loci (Supplementary Table 2).

We advanced these loci for replication in a GWAS of refractive error carried out by the UK Biobank Eye & Vision (UKEV) Consortium (*N*=95,505) [33] (Online Methods). Six out of the 167 variants were not considered for replication analysis. One of these five variants (rs3138141, *RDH5*) was identified previously and therefore still considered as a refractive error risk variant[26,27]. The remaining 161 genetic variants were tested for replication. 86% (138/161) of the candidate variants replicated significantly: 104 (65%) replicated surpassing genome-wide significance and 34 replicated surpassing Bonferroni *correction (P*<$3.0x10^{-4}$; 21.1%); another 12 showed nominal evidence for replication (0.05<*P*<$3.0x10^{-4}$; 7.5%) and only 11 (7%) did not replicate at all (Table 1, Supplementary Table 2).

As CREAM and 23andMe employed different phenotypic outcomes, we evaluated consistency of genotypic effects by comparing marker-wise additive genetic effect sizes (in units diopters per risk allele variant) for SphE from CREAM-EUR against those (in units log(HR) per risk allele variant) for AODM from 23andMe. All variants strongly associated with either outcome (*P*<0.001) were concordant in direction-of-effect, and had highly correlated effect sizes (Figure 2a,b; Supplementary Figure 8). For these variants a 10% decrease in log(HR) for AODM, indicating an earlier age-at-myopia onset, was associated with a decrease of 0.15 diopters in SphE. A quantitative analysis for all common SNPs (MAF>0.01; HapMap3) using LD score regression yielded a genetic correlation of 0.93 (95% CI 0.86-0.99; *P*=$2.1x10^{-159}$), confirming that effect sizes for both phenotypic outcomes were closely related.

Gene annotation of susceptibility loci

We annotated all genetic variants with wANNOVAR using the University of California Santa Cruz (UCSC) Known Gene database[34,35]. The identified 139 genetic loci were annotated to 208 genes and known transcribed RNA genes (Table 1, Supplementary Table 2, Online Methods). The physical positions of the lead genetic variants relative to protein-coding genes are shown in Figure 1c. 86% of the identified variants were either intragenic or less than 50 kb from the 5'or 3' end of the transcription start site. We found seven exonic variants (Supplementary Table 7) of which two had MAF≤0.05: rs5442 (*GNB3*) and rs17400325 (*PDE11A).* The index SNP in the *GNB3* locus with MAF 0.05 in Europeans is a highly conserved missense variant (G272S) predicted to be damaging by PolyPhen-2[36] and SIFT[37]. *PDE11A* is presumed to play a role in tumorigenesis, brain function, and inflammation[38]. The index SNP in the *PDE11A* locus with MAF 0.03 in Europeans is also a highly conserved missense variant (Y727C); this variant was predicted to be damaging by PolyPhen[36], SIFT[39] and align GVGD[40,41].The other exonic variants, rs1064583 (*COL10A1*), rs807037 (*KAZALD1*), rs1550094 (*PRSS56*), rs35337422 (*RD3L*), and rs6420484 (*TSPAN10*) were not predicted to be damaging.

The most significant variant (Stage 3; rs12193446, P=4.21x10^{-84}) resides on chromosome 6 within a non-coding RNA, *BC035400*, in an intron of the *LAMA2* gene. This locus had been identified previously, but our current fine mapping redefined the most associated variant. The function and potential downstream target sites for BC035400 are currently unknown. The previously most strongly associated variant, rs524952 on chromosome 15 near *GJD2*, was the second most significant variant (P=2.28x10^{-65}).

Post-GWAS analyses

We performed two gene-based tests, fastBAT[42] and EUGENE[43], and applied a functional enrichment approach using fgwas[44] (Online Methods). With fastBAT, we identified 13 genes at P <2.0x10^{-6}, one of which (*CHD7*) had been identified previously[26,27]. Using EUGENE, we found 7 genes at P <2.0x10^{-6} after incorporation of blood eQTLs. With fgwas, we identified 6 loci, which could be annotated to 9 genes, at posterior probability >0.9. Two genes (*HMGN4* and *TLX1*) showed significant associations in two or

more approaches. Taken together, these post-GWAS approaches resulted in a total of 22 additional candidate loci for refractive error, annotated to 25 genes (Supplementary Table 8). This increases the overall number of significant genetic associations to 161 candidate loci.

Polygenic risk scores

We calculated polygenic risk scores (PGRS)[45] per individual at various *P* thresholds (Online Methods) for Rotterdam Study I-III (RS I-III; *N*=10,792) after recalculating *P* and Z-scores of variants from Stage 3 excluding RS I-III. We found the highest fraction of phenotypic variance (7.8%) explained with 7,307 variants at *P* value threshold 0.005 (Supplementary Table 9). A PGRS based on these variants distinguished well between individuals with hyperopia and myopia at the lower and higher deciles (Figure 3); those in the highest decile had a 40-fold increased risk of myopia. When the PGRS was stratified for the median age (<63 or >63+ yrs), we found a significant difference in the variance explained (<63 yrs 8.9%; 63+ yrs 7.4%; *P* 0.0038). The variance explained by PGRS was not significantly different between males and females (8.3% vs 7.5%; *P* 0.13). The predictive value (area under the receiver operating characteristic curve, AUC) of the PGRS for myopia versus hyperopia adjusted for age and gender was 0.77 (95% CI=0.75–0.79), a 10% increase compared to previous estimations[46].

Trans-ethnic comparison of genotypic effects

To explore potential ancestry differences in the identified refractive error loci, we calculated the heritability explained by common genetic variants (SNP-h^2) for Europeans and Asians using LD score regression[47]. SNP-h^2 was 0.214 (95% CI 0.185- 0.243) and 0.172 (95% CI 0.154- 0.190) in the European samples (CREAM-EUR and 23andMe), while it was only 0.053 (95% CI -0.025- 0.131) in the Asian sample (CREAM-EAS). Next, we estimated the genetic correlation between Europeans and Asians by comparing variant effect size for common variants using Popcorn[48] (Online Methods). Two genetic correlation metrics were calculated T; First, a genetic effect correlation (ρge) that quantifies the

correlation in SNP effect sizes between Europeans and Asians without taking into account ancestry-related differences in allele frequency; and second, a genetic impact correlation (ρgi) that estimates the correlation in variance-normalized SNP effect sizes between the two ancestry groups (Table 2). Estimates of ρge were high between Europeans and Asians, but significantly different from 1 (0.79 and 0.80, respectively at $P < 1.9 \times 10^{-6}$; Table 2), indicating a clear genetic overlap but a difference in per allele effect size. Estimates of ρgi were similarly high (>0.8), but not significantly different from 1 for the correlation between CREAM-EUR and CREAM-ASN (P=0.065), indicating that the genetic impact of these alleles may still be similar.

In silico pathway analysis

We used an array of bioinformatics tools to investigate potential functions and pathways of the associated genes. We first employed DEPICT[49] to perform a gene set enrichment analysis, a tissue type enrichment, and a gene prioritization analysis, on all variants with $P < 1.00 \times 10^{-5}$ from Stage 3. The gene set enrichment analysis resulted in 66 reconstituted gene sets, of which 55 (83%) were eye-related. To reduce redundancies between pathways, we clustered the significant pathways into 13 meta gene sets (false discovery rate (FDR) <5% and a P <0.05) (Supplementary Note, Figure 4, Supplementary Table 10). The most significant gene set was the 'abnormal photoreceptor inner segment morphology' (MP:0003730; $P = 1.79 \times 10^{-7}$). The eye-related meta gene sets consisted of the 'thin retinal outer nuclear layer' (MP:0008515; 27 (55%) gene sets), 'detection of light stimulus' (GO:0009583; 13 (24%) gene sets), 'nonmotile primary cilium' (GO:0031513; 4 (6%) gene sets), and 'abnormal anterior eye segment morphology' (MP:0005193; 4 (6%) gene sets). The first three meta gene sets had a Pearson's correlation > 0.6. Interestingly, *RGR*, *RP1L1*, *RORB* and *GNB3* were present in all of these meta gene sets. Retina was the most significant tissue of expression according to the tissue enrichment analysis ($P = 1.11 \times 10^{-4}$, FDR <0.01). From the gene prioritization according to DEPICT, 7 genes were highlighted as the most likely causal genes at $P < 7.62 \times 10^{-6}$ and FDR<0.05: *ANO2*, *RP1L1*, *GNB3*, *EDN2*, *RORB* and *CABP4*.

Next, we performed a canonical pathway analysis on all genes annotated to the variants of Stage 3 using Ingenuity Pathway Analysis (See URLs). All genes were run against the IPA database incorporating functional biological evidence on genomic and proteomic expression based on regulation or binding studies. IPA identified "Glutamate Receptor Signaling" with central player *Nf-kB* gene as the most significant pathway after correction for multiple testing (ratio of the number of molecules 8.8% and Fisher's Exact test P=1.56x10^{-4}; Supplementary Figure 9).

From disease-associated loci to biological mechanisms

We adapted the scoring scheme designed by Fritsche et al.[50] to highlight genes for which there is biological plausibility for a role in eye growth[50]. We used 10 equally rated categories (Online Methods; Figure 5; Supplementary Table 11; Supplementary Note). One-hundred-and-nine index variants replicated in two or more individual cohorts; we found evidence for seven genetic variants with eQTL effects in multiple tissue types; nine exonic variants, of which seven predicted protein-alterations (Supplementary Table 7); 31 RNA genes, five located in the 3' or 5'UTR (Supplementary Table 12, Supplementary Figure 10), 84 genes resulting in an ocular phenotype in humans (Supplementary Table 13) and 36 in mice (Supplementary Table 14); 172/212 (81%) genes expressed in human ocular tissue (Supplementary Note, Supplementary Table 15); 41 genes identified by DEPICT at P <5.4x10^{-4} and FDR<0.05 and 45 genes contributed to the most significant canonical pathways of IPA. Notably, 48 of the associated genes encode known drug targets (Supplementary Table 16).

The gene with the highest biological plausibility score (score=8) was *GNB3*, a highly conserved gene encoding a guanine nucleotide-binding protein expressed in rod and cone photoreceptors and ON-bipolar cells[51]. *GNB3* participates in signal transduction through G-protein-coupled receptors and enhances the temporal accuracy of phototransduction and ON-center signaling in the retina[51]. As described above, the index SNP harbors a missense variant associated with refractive errors. Non-synonymous mutations within *GNB3* are known to cause syndromic congenital stationary night

blindness[52] in humans, progressive retinopathy and globe enlargement in chickens[51], and abnormal development of the photoreceptor-bipolar synapse in knock-out mice[53,54].

Other genes highly ranked (score=7) include *CYP26A1*, *GRIA4*, *RDH5*, *RORB*, and *RGR*, all previously associated with refractive error, and one newly identified gene: *EFEMP1*. *EFEMP1* encodes a member of the fibulin family of extracellular matrix glycoproteins, and is found pan-ocularly including in the inner nuclear layer and Bruch's membrane. Mutations in this gene lead to specific macular dystrophies[55], while variants have also been shown to co-segregate with primary open-angle glaucoma[56] and associate with optic disc cup area[57].

Several other genes are noteworthy for their function. *CABP4*, a calcium-binding protein expressed in cone and rod photoreceptor cells, mediates Ca^{2+}-influx and glutamate release in the photoreceptor-bipolar synapse[58]. Mutations in this gene have been described in congenital cone-rod synaptic disorder[59], a retinal dystrophy associated with nystagmus, photophobia, and, remarkably, high hyperopia. *KCNMA1* encodes pore-forming alpha subunits of Ca^{2+}-activated K^+ (BK) channels. These channels regulate synaptic transmission exclusively in the rod pathway[60]. *ANO2* is a Ca^{2+}-activated Cl^- channel recently reported to regulate retinal pigment epithelial (RPE) cell volume in a light-dependent manner[64]. *EDN2* is a potent vasoconstrictor that binds to two G-protein-coupled receptors, *EDNRA*, which resides on bipolar dendrites, and *EDNRB*, which is present on Mueller and horizontal cells. Both receptors are also present on choroidal vessels[65], implying that the choroid as well as retinal cells are target sites for this gene. *RP1L1* is expressed in cone and rod photoreceptors where it is involved in the maintenance of microtubules in the connecting cilium[66]. Mutations in this gene cause dominant macular dystrophy and retinitis pigmentosa[67]. We replicated two genes known to cause myopia in family studies. *FBN1* harbors mutations causing with Marfan (OMIM #154700) and Weil Marchesani (OMIM #608328) syndrome; *PTPRR* was one of the candidates in the MYP3 locus, which was found by linkage in families with high myopia[68].

The location of rs7449443 (P=3.58x10^{-8}) is notable as it resides in between *DRD1* and *FLJ16171*. *DRD1* encodes dopamine receptor 1 and is known to modulate dopamine receptor 2-mediated events[69,70].

The dopamine pathway has been implicated in myopia pathogenesis in many studies[69,71]. SNPs in and near other genes involved in the dopamine pathway (dopamine receptors, synthesis, degradation, and transporters)[72-74] did not reveal genome-wide significant associations (Supplementary Note, Supplementary Table 17; Supplementary Figure 11).

There were 31 genetic variants in or near DNA structures transcribing RNA genes (non coding RNA, linc RNAs, tRNAs, snoRNas, rRNAs). Notably, five were in the transcription region and thirteen were in the vicinity (>0 kb and ≤50 kb) of start or end of the RNA gene transcription region. They received low scores, since many have no reported function or disease association to date (Figure 5, Supplementary Figure 10, Supplementary Table 12). Our ranking of genes based on functional information existing in the public domain does not necessarily represent the true order of importance for refractive error pathogenesis. The observation that genes with strong statistical association were distributed over all scores supports this concept. Nevertheless, this list may help to select genes for subsequent functional studies.

Finally, integration of all our findings, supported by literature, allowed us to annotate a large number of genes to ocular cell types (Figure 6). Remarkably, all cell types of the retina harbored refractive error genes, as well as the RPE, vascular endothelium, and extracellular matrix.

Genetic pleiotropy

We performed a GWAS catalogue look up using FUMA to investigate overlap of genes with other common traits (Supplementary Figure 12)[75]. Refractive error and hyperopia were replicated significantly after correcting for multiple testing (adjusted *P* value=1.44×10^{-52} and 9.34×10^{-9}, respectively). We found significant overlap with 74 other traits, of which height (adjusted *P* value=1.11×10^{-10}), obesity (adjusted *P*=1.38×10^{-10}), and BMI (adjusted *P*=4.05×10^{-7}) were most important. Ocular diseases significantly associated were glaucoma (optic cup area, intraocular pressure; adjusted *P*=2.69×10^{-5} and 3.01×10^{-5}, respectively) and age-related macular degeneration (adjusted *P* value=1.27×10^{-3}).

DISCUSSION

Myopia may become the leading cause of world blindness in the near future, a grim outlook for which current counteractions are still insufficient[11,76]. To improve understanding of the genetic landscape and biology of refractive error, we conducted a large GWAS meta-analysis in 160,420 participants of mixed ancestry and replicated in 95,505 participants. This led to the identification of 139 independent susceptibility loci by single variant analysis and 22 additional loci through post-GWAS methods, a four-fold increase in refractive error genes. The majority of annotated genes were found to be expressed in the human posterior segment of the eye. Using in silico analysis, we identified significant biological pathways, of which retinal cell physiology, light processing, and, specifically, glutamate receptor signaling were the most prominent mechanisms. Our integrated bio-informatic approach highlighted known ocular functionality for many genes.

To ensure robustness of our genetic associations, we included studies of various designs and populations, sought replication in an independent cohort of significant sample size, and stringently accounted for population stratification by performing genomic control at all stages of the meta-analysis[77]. We combined studies with outcomes based on actual refractive error measurements as well as on self-reported age-of-myopia-onset, and found the direction-of-effect of the associated variants, as well as their effect size, to be remarkably consistent. Combining two different outcome measures may appear unconventional, but age of onset and refractive error have been shown to be very tightly correlated[11,28,78,79]. [78,79]Moreover, the high genetic correlation (93%) of common SNPs between the two phenotypes underscores their similarity. Most compelling evidence was provided by replication of 86% of the discovered variants in the independent UKEV which also used conventional refractive error measurements. This robustness indicates that both phenotypic outcomes can be used to capture a shared source of genetic variation. In addition, we found trans-ethnic replication of significant loci, and a high correlation of genetic effects of common variants in the Europeans and Asians. Our findings support a

largely shared genetic predisposition to refractive error and myopia in the two ethnicities, although ancestry-specific allelic effects may exist. The low heritability estimate in Asians may, in part, be explained by the low representation of this ethnicity in our study sample. Alternatively, it may imply that environmental factors explain a greater proportion of the phenotypic risk and recent rise in myopia prevalence in this ancestry group[80].

Limitations of our study were the possibility of false negative findings due to genomic control, and underrepresentation of studies with Asian ancestry. Heterogeneity of observed effect estimates was large for several associated variants, but not unexpected, given the large number of collaborating studies with varying methodology.

Although neurotransmission was previously suggested pathway[26,27], our current pathway analyses provide more in-depth insights into the retinal circuitry driving refractive error. DEPICT identified 'thin retinal outer nuclear layer', 'detection of light stimulus', and 'nonmotile primary cilium' as the most important meta-gene sets. These are the main characteristics of photoreceptors, which are located in the outer retina and contain cilia. These photosensitive cells drive the phototransduction cascade in response to light, which in turn induces visual information processing. IPA pointed towards glutamate receptor signaling as the most significant pathway. Glutamate is released by photoreceptors and determines conductance of retinal signaling to the ON and OFF bipolar cells[81]. Our functional gene look ups provide evidence that rod (*CLU*) as well as cone (*GNB3*) bipolar cells play a role. Taken together, these findings strongly suggest that light response and light processing in the retina are initiating factors leading to refractive error.

The genetic association with light-dependent pathways may also link to the well-established protective effect of outdoor exposure on myopia. We found suggestive evidence for a genetic association with *DRD1*. The dopaminergic pathway has been studied extensively in animal models for its role in controlling eye growth in response to light[69,71,82-91]. *DRD1* was found to be a mediator in this process, as bright light increased DRD1 activity in the bipolar ON-pathway, and diminished form-deprivation myopia in mice. Blockage of DRD1 reversed this inhibitory effect[92]. We did not find evidence for direct

involvement of other genes in the dopamine pathway, but *GNB3* may be an indirect modifier as it is a downstream signaling molecule of dopamine and has been shown to influence availability of the dopamine transporter DAT[93]. Although a promising target for therapy, further evidence of *DRD1* in human myopiagenesis is warranted.

Novel pathways elicited by the newly identified genes are anterior segment morphology *(TCF7L2, VIPR2, MAF*) and angiogenesis (*FLT1*). In addition, the high number of variants residing near small RNA genes suggests that post-transcriptional regulation is an important mechanism, as these RNAs are known to play a distinct and central regulatory role in cells[94]. These findings will serve as leads for future studies performing detailed mapping of cellular networks, and functional studies on genes implicated in ocular phenotypes, harboring protein-altering variants, and proven drug targets.

Our evaluation of shared genetics between refractive error and other disease-relevant phenotypes highlighted overlap with anthropometric traits such as height, obesity, and body mass index. This could give valuable additional clues as to the phenotypic outcomes of perturbations of some of the networks identified.

Our genetic observations add credence to the current notion that refractive errors are caused by a retina-to-sclera signaling cascade that induces scleral remodeling in response to light stimuli. The concept of this cascade originates from various animal models showing that form deprivation, retinal defocus and contrast, ambient light, and wavelength can influence eye growth in young animals[95-97]. Cell-specific moieties in this putative signaling cascade in humans were largely unknown, although animal models implicated GABA, dopamine, all-trans-retinoic acid and TGF-β[69,91,98,99]. Our study provides a large number of new molecular candidates for this cascade, and clearly shows that a wide range of neuronal cell types in the retina, the RPE, the vascular endothelium, as well as components of the extracellular matrix are implicated. The many interprotein relationships exemplify the complexity of eye growth, and provide a challenge to develop strategies to prevent pathological eye elongation.

In conclusion, by using a cross-ancestry design in the largest study population on common refractive errors to date, we uncovered numerous novel loci and pathways involved in eye growth. Our

multi-disciplinary approach incorporating GWAS data with *in silico* analyses and expression experiments provides an example for the design of future genetic studies for complex traits. Additional genetic insights into refractive errors will be gained by increasing sample size and genotyping depth, by performing family studies to identify rare alleles of large effects, and by evaluating population extremes. Our list of plausible genes and pathways provide a plethora of data for future studies focusing on gene-environment interaction, and on translation of GWAS findings into starting points for therapy.

URLs. LDSC https://github.com/bulik/ldsc; Popcorn https://github.com/brielin/Popcorn; OMIM http://omim.org; wANNOVAR http://wannovar.wglab.org/; Polyphen http://genetics.bwh.harvard.edu/pph2/; SIFT http://sift.jcvi.org/www/SIFT_aligned_seqs_submit.html; Mutation Taster http://www.mutationtaster.org/; IPA http://www.ingenuity.com/index.html;

Acknowledgments

We gratefully thank the invaluable contributions of all study participants, their relatives and staff at the recruitment centers. We thank all contributors to the CREAM Consortium, 23andMe, and UKEV for their generosity in sharing data and help in the production of this publication. Funding for this particular GWAS mega-analysis was provided by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant 648268), Netherlands Organisation for Scientific Research (NWO, grant 91815655) and the National Eye Institute (grant R01EY020483). Funding agencies which facilitated the execution of the individual studies are acknowledged in the **Supplementary Note**.

Competing interests

N.A.F., N.E., J.Y.T., and the 23andMe Research Team are current or former employees of 23andMe, Inc., and hold stock or stock options in 23andMe. J.B. J. is a patent holder with Biocompatibles UK Ltd. (Franham, Surrey, UK) (Title: Treatment of eye diseases using encapsulated cells encoding and secreting

neuroprotective factor and / or anti-angiogenic factor; Patent number: 20120263794), and patent application with University of Heidelberg (Heidelberg, Germany) (Title: Agents for use in the therapeutic or prophylactic treatment of myopia or hyperopia; European Patent Number: 3 070 101). The other authors declare no competing financial interests.

Author contributions

M.S.T., V.J.M.V., S.M., J.A.G., A.I.I., R.W., P.G.H., A.I.I., and E.M.v.L. performed the analyses. C.C.W.K., V.J.M.V., M.S.T., R.W., J.A.G., and S.M. drafted the manuscript, and C.J.H., P.G.H., A.P.K., C.M.v.D., D.S, E.M.v.L., J.E.B.W., J.Y.T., N.A.F., Q.F., S.M.S., and V.V. critically reviewed the manuscript. A.N., A.P.K., A.T., C.B., C.Gi., C.L.S., C.Y.C., G.Bi., G.C., I.R., J.E.B.W., J.E.H., J.S.Ri., J.W., J.X., K.M.W., K.Y., P.M.C., S.M.H., M.S.T., N.A.F., N.E., P.C., P.Gh., P.K.J., Q.F., R.Ho., R.L.S., R.P.I., R.W., T.H., T.H.S.A., T.Z., V.V., W.Y.S., W.Z., X.L.S., Y.C.T., Y.S., and Y.Y.T. performed data analysis for the individual studies; and A.D.P., A.G.U., A.T., A.W.H., B.E.K.K., C.C.W.K., C.D., C.Gr., C.H., C.J.H., C.W., C.Y.C., D.A.M., F.R., G.Be., H.M.H., J.A.G., J.B.J., J.E.B.W., J.E.C., J.F.W., J.H.L., J.R.V., J.S.Ra., J.S.Ri., J.Y.T., K.Y., M.A.M.S., N.G.M., N.P., O.Po., O.Pa., O.T.R., P.Gu., P.J.F., P.M., P.N.B., R.K., S.K.I., S.M.S., T.L., T.M., W.Z., Y.C.T., and Y.X.W. contributed to data assembly. A.A.B.B., A.W., C.Gr., D.S., K.N.W., S.W.T., and T.L.Y. performed expression experiments, and M.S.T., A.A.B.B., P.J.v.d.S., and R.Ha. performed in silico pathway analyses. C.C.W.K. and C.J.H. conceived and designed the outline of the current report, and jointly with A.M., A.H., A.W.H., C.D., C.H., C.J.H., C.M.v.D., C.W., C.Y.C., D.A.M., D.S., E.S.T., F.M., G.Bi., I.R., J.A.G., J.B.J., J.E.B.W., J.E.C., J.F.W., J.H.L., J.R.V., J.Y.T., N.A., N.A.F., N.P., O.Pa., O.T.R., P.J.F., P.N.B., S.K.I., S.M.S., T.L., T.Y.W., T.L.Y., V.V., Y.X.W., and Y.Y.T. supervised conduction of experiments and analyses.

References for main text

1. Pan, C.W., Ramamurthy, D. & Saw, S.M. Worldwide prevalence and risk factors for myopia. *Ophthalmic Physiol Opt* **32**, 3-16 (2012).
2. Morgan, I.G. What Public Policies Should Be Developed to Deal with the Epidemic of Myopia? *Optom Vis Sci* **93**, 1058-60 (2016).
3. Morgan, I. & Rose, K. How genetic is school myopia? *Prog Retin Eye Res* **24**, 1-38 (2005).

4. Morgan, I.G., Ohno-Matsui, K. & Saw, S.M. Myopia. *Lancet* **379**, 1739-48 (2012).
5. Williams, K.M. *et al.* Increasing Prevalence of Myopia in Europe and the Impact of Education. *Ophthalmology* **122**, 1489-97 (2015).
6. Williams, K.M. *et al.* Prevalence of refractive error in Europe: the European Eye Epidemiology (E(3)) Consortium. *Eur J Epidemiol* **30**, 305-15 (2015).
7. Vongphanit, J., Mitchell, P. & Wang, J.J. Prevalence and progression of myopic retinopathy in an older population. *Ophthalmology* **109**, 704-11 (2002).
8. Seet, B. *et al.* Myopia in Singapore: taking a public health approach. *Br J Ophthalmol* **85**, 521-6 (2001).
9. Smith, T.S., Frick, K.D., Holden, B.A., Fricke, T.R. & Naidoo, K.S. Potential lost productivity resulting from the global burden of uncorrected refractive error. *Bull World Health Organ* **87**, 431-7 (2009).
10. Verhoeven, V.J. *et al.* Visual consequences of refractive errors in the general population. *Ophthalmology* **122**, 101-9 (2015).
11. Tideman, J.W. *et al.* Association of Axial Length With Risk of Uncorrectable Visual Impairment for Europeans With Myopia. *JAMA Ophthalmol* **134**, 1355-1363 (2016).
12. Flitcroft, D.I. The complex interactions of retinal, optical and environmental factors in myopia aetiology. *Prog Retin Eye Res* **31**, 622-60 (2012).
13. Nakanishi, H. *et al.* A genome-wide association analysis identified a novel susceptible locus for pathological myopia at 11q24.1. *PLoS Genet* **5**, e1000660 (2009).
14. Lam, C.Y. *et al.* A genome-wide scan maps a novel high myopia locus to 5p15. *Invest Ophthalmol Vis Sci* **49**, 3768-78 (2008).
15. Stambolian, D. *et al.* Meta-analysis of genome-wide association studies in five cohorts reveals common variants in RBFOX1, a regulator of tissue-specific splicing, associated with refractive error. *Hum Mol Genet* **22**, 2754-64 (2013).
16. Fan, Q. *et al.* Genetic variants on chromosome 1q41 influence ocular axial length and high myopia. *PLoS Genet* **8**, e1002753 (2012).
17. Fan, Q. *et al.* Meta-analysis of gene-environment-wide association scans accounting for education level identifies additional loci for refractive error. *Nat Commun* **7**, 11008 (2016).
18. Cheng, C.Y. *et al.* Nine Loci for Ocular Axial Length Identified through Genome-wide Association Studies, Including Shared Loci with Refractive Error. *Am J Hum Genet* **93**, 264-77 (2013).
19. Shi, Y. *et al.* Exome sequencing identifies ZNF644 mutations in high myopia. *PLoS Genet* **7**, e1002084 (2011).
20. Shi, Y. *et al.* Genetic variants at 13q12.12 are associated with high myopia in the Han Chinese population. *Am J Hum Genet* **88**, 805-13 (2011).
21. Li, Y.J. *et al.* Genome-wide association studies reveal genetic variants in CTNND2 for high myopia in Singapore Chinese. *Ophthalmology* **118**, 368-75 (2011).
22. Li, Z. *et al.* A genome-wide association study reveals association between common variants in an intergenic region of 4q25 and high-grade myopia in the Chinese Han population. *Hum Mol Genet* **20**, 2861-8 (2011).
23. Liu, J. & Zhang, H.X. Polymorphism in the 11q24.1 genomic region is associated with myopia: a comprehensive genetic study in Chinese and Japanese populations. *Mol Vis* **20**, 352-8 (2014).
24. Tran-Viet, K.N. *et al.* Mutations in SCO2 are associated with autosomal-dominant high-grade myopia. *Am J Hum Genet* **92**, 820-6 (2013).
25. Aldahmesh, M.A. *et al.* Mutations in LRPAP1 are associated with severe myopia in humans. *Am J Hum Genet* **93**, 313-20 (2013).
26. Verhoeven, V.J. *et al.* Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat Genet* (2013).
27. Kiefer, A.K. *et al.* Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet* **9**, e1003299 (2013).

28. Wojciechowski, R. & Hysi, P.G. Focusing in on the complex genetics of myopia. *PLoS Genet* **9**, e1003442 (2013).
29. Genomes Project, C. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
30. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
31. Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* **19**, 807-12 (2011).
32. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. FUMA: Functional mapping and annotation of genetic associations. *bioRxiv* (2017).
33. Plotnikov, D., Guggenheim, J. & Is a large eye size a risk factor for myopia? A Mendelian randomization study. *bioRxiv* (2017).
34. UCSC Genome Browser.
35. Hsu, F. *et al.* The UCSC Known Genes. *Bioinformatics* **22**, 1036-46 (2006).
36. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-9 (2010).
37. Ng, P.C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res* **31**, 3812-4 (2003).
38. Kelly, M.P. Does phosphodiesterase 11A (PDE11A) hold promise as a future therapeutic target? *Curr Pharm Des* **21**, 389-416 (2015).
39. Kumar, P., Henikoff, S. & Ng, P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4**, 1073-81 (2009).
40. Mathe, E. *et al.* Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods. *Nucleic Acids Res* **34**, 1317-25 (2006).
41. Tavtigian, S.V. *et al.* Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* **43**, 295-305 (2006).
42. Bakshi, A. *et al.* Fast set-based association analysis using summary data from GWAS identifies novel gene loci for human complex traits. *Sci Rep* **6**, 32894 (2016).
43. Ferreira, M.A. *et al.* Gene-based analysis of regulatory variants identifies 4 putative novel asthma risk genes related to nucleotide synthesis and signaling. *J Allergy Clin Immunol* (2016).
44. Pickrell, J.K. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am J Hum Genet* **94**, 559-73 (2014).
45. International Schizophrenia, C. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748-52 (2009).
46. Verhoeven, V.J. *et al.* Large scale international replication and meta-analysis study confirms association of the 15q14 locus with myopia. The CREAM consortium. *Hum Genet* **131**, 1467-80 (2012).
47. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
48. Brown, B.C., Asian Genetic Epidemiology Network Type 2 Diabetes, C., Ye, C.J., Price, A.L. & Zaitlen, N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *Am J Hum Genet* **99**, 76-88 (2016).
49. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
50. Fritsche, L.G. *et al.* A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat Genet* **48**, 134-43 (2016).
51. Ritchey, E.R. *et al.* Vision-guided ocular growth in a mutant chicken model with diminished visual acuity. *Exp Eye Res* **102**, 59-69 (2012).
52. Vincent, A. *et al.* Biallelic Mutations in GNB3 Cause a Unique Form of Autosomal-Recessive Congenital Stationary Night Blindness. *Am J Hum Genet* **98**, 1011-9 (2016).

53. Blake, J.A. *et al.* Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res* **45**, D723-D729 (2017).
54. Nikonov, S.S. *et al.* Cones respond to light in the absence of transducin beta subunit. *J Neurosci* **33**, 5182-94 (2013).
55. Stone, E.M. *et al.* A single EFEMP1 mutation associated with both Malattia Leventinese and Doyne honeycomb retinal dystrophy. *Nat Genet* **22**, 199-202 (1999).
56. Mackay, D.S., Bennett, T.M. & Shiels, A. Exome Sequencing Identifies a Missense Variant in EFEMP1 Co-Segregating in a Family with Autosomal Dominant Primary Open-Angle Glaucoma. *PLoS One* **10**, e0132529 (2015).
57. Springelkamp, H. *et al.* ARHGEF12 influences the risk of glaucoma by increasing intraocular pressure. *Hum Mol Genet* **24**, 2689-99 (2015).
58. Haeseleer, F. *et al.* Essential role of Ca2+-binding protein 4, a Cav1.4 channel regulator, in photoreceptor synaptic function. *Nat Neurosci* **7**, 1079-87 (2004).
59. Littink, K.W. *et al.* A novel homozygous nonsense mutation in CABP4 causes congenital cone-rod synaptic disorder. *Invest Ophthalmol Vis Sci* **50**, 2344-50 (2009).
60. Grimes, W.N., Li, W., Chavez, A.E. & Diamond, J.S. BK channels modulate pre- and postsynaptic signaling at reciprocal synapses in retina. *Nat Neurosci* **12**, 585-92 (2009).
61. Parker, R.O. & Crouch, R.K. Retinol dehydrogenases (RDHs) in the visual cycle. *Exp Eye Res* **91**, 788-92 (2010).
62. Radu, R.A. *et al.* Retinal pigment epithelium-retinal G protein receptor-opsin mediates light-dependent translocation of all-trans-retinyl esters for synthesis of visual chromophore in retinal pigment epithelial cells. *J Biol Chem* **283**, 19730-8 (2008).
63. Luo, T., Sakai, Y., Wagner, E. & Drager, U.C. Retinoids, eye development, and maturation of visual function. *J Neurobiol* **66**, 677-86 (2006).
64. Keckeis, S., Reichhart, N., Roubeix, C. & Strauss, O. Anoctamin2 (TMEM16B) forms the Ca2+-activated Cl- channel in the retinal pigment epithelium. *Exp Eye Res* **154**, 139-150 (2016).
65. Prasanna, G., Narayan, S., Krishnamoorthy, R.R. & Yorio, T. Eyeing endothelins: a cellular perspective. *Mol Cell Biochem* **253**, 71-88 (2003).
66. Yamashita, T. *et al.* Essential and synergistic roles of RP1 and RP1L1 in rod photoreceptor axoneme and retinitis pigmentosa. *J Neurosci* **29**, 9748-60 (2009).
67. Davidson, A.E. *et al.* RP1L1 variants are associated with a spectrum of inherited retinal diseases including retinitis pigmentosa and occult macular dystrophy. *Hum Mutat* **34**, 506-14 (2013).
68. Hawthorne, F. *et al.* Association mapping of the high-grade myopia MYP3 locus reveals novel candidates UHRF1BP1L, PTPRR, and PPFIA2. *Invest Ophthalmol Vis Sci* **54**, 2076-86 (2013).
69. Feldkaemper, M. & Schaeffel, F. An updated view on the role of dopamine in myopia. *Exp Eye Res* **114**, 106-19 (2013).
70. Paul, M.L., Graybiel, A.M., David, J.C. & Robertson, H.A. D1-like and D2-like dopamine receptors synergistically activate rotation and c-fos expression in the dopamine-depleted striatum in a rat model of Parkinson's disease. *J Neurosci* **12**, 3729-42 (1992).
71. Stone, R.A., Lin, T., Laties, A.M. & Iuvone, P.M. Retinal dopamine and form-deprivation myopia. *Proc Natl Acad Sci U S A* **86**, 704-6 (1989).
72. Gardner, M., Bertranpetit, J. & Comas, D. Worldwide genetic variation in dopamine and serotonin pathway genes: implications for association studies. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 1070-5 (2008).
73. D'Souza, U.M. & Craig, I.W. Functional polymorphisms in dopamine and serotonin pathway genes. *Hum Mutat* **27**, 1-13 (2006).
74. Beaulieu, J.M. & Gainetdinov, R.R. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacol Rev* **63**, 182-217 (2011).
75. MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901 (2017).

76. Holden, B.A. *et al.* Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. *Ophthalmology* **123**, 1036-42 (2016).
77. Cardon, L.R. & Palmer, L.J. Population stratification and spurious allelic association. *Lancet* **361**, 598-604 (2003).
78. Chua, S.Y. *et al.* Age of onset of myopia predicts risk of high myopia in later childhood in myopic Singapore children. *Ophthalmic Physiol Opt* **36**, 388-94 (2016).
79. Williams, K.M. *et al.* Age of myopia onset in a British population-based twin cohort. *Ophthalmic Physiol Opt* **33**, 339-45 (2013).
80. Dolgin, E. The myopia boom. *Nature* **519**, 276-8 (2015).
81. Connaughton, V. Glutamate and Glutamate Receptors in the Vertebrate Retina. in *Webvision: The Organization of the Retina and Visual System* (eds. Kolb, H., Fernandez, E. & Nelson, R.) (Salt Lake City (UT), 1995).
82. Hung, G.K., Mahadas, K. & Mohammad, F. Eye growth and myopia development: Unifying theory and Matlab model. *Comput Biol Med* **70**, 106-18 (2016).
83. Norton, T.T. What Do Animal Studies Tell Us about the Mechanism of Myopia-Protection by Light? *Optom Vis Sci* **93**, 1049-51 (2016).
84. Weiss, S. & Schaeffel, F. Diurnal growth rhythms in the chicken eye: relation to myopia development and retinal dopamine levels. *J Comp Physiol A* **172**, 263-70 (1993).
85. Stone, R.A., Lin, T., Iuvone, P.M. & Laties, A.M. Postnatal control of ocular growth: dopaminergic mechanisms. *Ciba Found Symp* **155**, 45-57; discussion 57-62 (1990).
86. Morgan, I.G. The biological basis of myopic refractive error. *Clin Exp Optom* **86**, 276-88 (2003).
87. Li, X.X., Schaeffel, F., Kohler, K. & Zrenner, E. Dose-dependent effects of 6-hydroxy dopamine on deprivation myopia, electroretinograms, and dopaminergic amacrine cells in chickens. *Vis Neurosci* **9**, 483-92 (1992).
88. Iuvone, P.M., Tigges, M., Stone, R.A., Lambert, S. & Laties, A.M. Effects of apomorphine, a dopamine receptor agonist, on ocular refraction and axial elongation in a primate model of myopia. *Invest Ophthalmol Vis Sci* **32**, 1674-7 (1991).
89. Ashby, R., McCarthy, C.S., Maleszka, R., Megaw, P. & Morgan, I.G. A muscarinic cholinergic antagonist and a dopamine agonist rapidly increase ZENK mRNA expression in the form-deprived chicken retina. *Exp Eye Res* **85**, 15-22 (2007).
90. Ashby, R. Animal Studies and the Mechanism of Myopia-Protection by Light? *Optom Vis Sci* **93**, 1052-4 (2016).
91. Rymer, J. & Wildsoet, C.F. The role of the retinal pigment epithelium in eye growth regulation and myopia: a review. *Vis Neurosci* **22**, 251-61 (2005).
92. Chen, S. *et al.* Bright Light Suppresses Form-Deprivation Myopia Development With Activation of Dopamine D1 Receptor Signaling in the ON Pathway in Retina. *Invest Ophthalmol Vis Sci* **58**, 2306-2316 (2017).
93. Chen, P.S. *et al.* Effects of C825T polymorphism of the GNB3 gene on availability of dopamine transporter in healthy volunteers--a SPECT study. *Neuroimage* **56**, 1526-30 (2011).
94. Scott, M.S. & Ono, M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie* **93**, 1987-92 (2011).
95. McFadden, S.A. Understanding and Treating Myopia: What More We Need to Know and Future Research Priorities. *Optom Vis Sci* **93**, 1061-3 (2016).
96. Smith, E.L., 3rd, Hung, L.F. & Arumugam, B. Visual regulation of refractive development: insights from animal studies. *Eye (Lond)* **28**, 180-8 (2014).
97. Zhang, Y. & Wildsoet, C.F. RPE and Choroid Mechanisms Underlying Ocular Growth and Myopia. *Prog Mol Biol Transl Sci* **134**, 221-40 (2015).
98. Harper, A.R. & Summers, J.A. The dynamic sclera: extracellular matrix remodeling in normal ocular growth and myopia development. *Exp Eye Res* **133**, 100-11 (2015).
99. Summers, J.A. The choroid as a sclera growth regulator. *Exp Eye Res* (2013).

Figure legends

Figure 1. GWAS meta-analysis identifies 140 loci for refractive error (Stage 3)

(a) We conducted a meta-analysis of genome-wide single-variant analyses for >10 million variants in 160,420 participants of CREAM and 23andMe (Stage 3). Shown is the Manhattan plot depicting *P* for association, highlighting new ($P < 5 \times 10^{-8}$ for the first time; green) and known (dark grey) refractive error loci previously found using HapMap II imputations from Kiefer et al.[27] and Verhoeven et al.[26] (**Table 1**). The horizontal lines indicate suggestive significance ($P=1\times10^{-5}$) or genome-wide significance ($P=5\times10^{-8}$). **(b)** We compared the minor allele frequencies of the 140 discovered index variants based on 1000G (blue: Europeans; red: Asians) to the minor allele frequencies of the previously found genetic variants based on HapMap II (green: Europeans; purple: Asians). Observed are an increase in genetic variants found across all minor allele frequency bins increase, including the lower minor allele frequency bins. **(c)** We annotated the 167 loci to genes using wANNOVAR. Shown are the distances between index variants from the nearest gene and its gene on the 5' and/or 3' site. The majority of index variants (84%) were at a distance of less than 50 kb up- or downstream from the annotated gene.

Figure 2. Correlation of statistical significance and effect size of SNPs based on spherical equivalent (SphE) in diopters and age of diagnosis of myopia (AODM) in years.

(a) *P* comparison of all genetic variants with $P < 1.0 \times 10^{-3}$ (n=7249) between CREAM meta-analysis (Stage 1) and 23andMe (Stage 2) meta-analysis. Shown is the overlap (red) and the difference (green) in *P* signals per cohort for genetic variants. Green genetic variants are only genome wide significant in either CREAM or 23andMe. Blue: genetic variants with P between 5.0×10^{-8} and 1.0×10^{-3} in both CREAM and 23andMe. **(b)** Comparison of effects (SphE and logHR of AODM in years; $P < 1.0 \times 10^{-3}$; n=7249) between CREAM and 23andMe. Same color code was applied as in (a). The effects were

concordant in their direction of effect on refractive error. We performed a simple linear regression between the effects of CREAM and 23andMe; the regression slope is -0.15 diopters per logHR of AODM in years.

Figure 3. Risk of refractive error per decile of polygenic risk score (Rotterdam Study I-III, N=10,792)

Distribution of refractive error in subjects from Rotterdam Study I–III (*N*=10,792) as a function of the optimal polygenic risk score (including 7,303 variants at $P \leq 0.005$ explaining 7.8% of the variance of SphE; Supplementary Table 9). Mean OR of myopia (black line) was calculated per polygenic risk score category using the lowest category as a reference. High myopia (SphE ≤-6 D), moderate myopia (SphE >-6 D & ≤ −3 D), low myopia (SphE > −3 D & <-1.5 D), emmetropia (SphE ≥ −1.5 D and ≤ 1.5 D), low hyperopia (SphE > 1.5 D & < 3 D), moderate hyperopia (SphE ≥ 3 D & 6 D), high hyperopia (SphE ≥ 6 D).

Figure 4. Visualization of the DEPICT gene-set enrichment analysis based on loci associated with refractive error and the correlation between the (meta)gene sets

(a) Shown are the 66 significantly enriched reconstituted gene sets clustered into thirteen meta gene sets based on the gene set enrichment analysis of DEPICT (pairwise Pearson correlations; $P < 0.05$). All genetic variants with a $P < 1 \times 10{-}5$ in the GWAS meta-analysis of stage 3 (n=21,073) and an FDR < 0.05 were considered. **(b)** Visualization of the interconnectivity between gene sets (n=13; pairwise Pearson correlations; $P < 0.05$) of the meta gene set 'Detection of Light Stimulus' (GO:0009583). **(c)** Visualization of the interconnectivity between gene sets (n=27; pairwise Pearson correlations; $P < 0.05$) of the largest meta gene set 'Thin Retinal Outer Nuclear Layer' (MP:0008515). In all panels, (meta)gene sets are represented by nodes colored according to statistical significance, and similarities between them are indicated by edges scaled according to their correlation; Pearson's $r \geq 0.2$ are shown in panel **(a)** and Pearson's $r \geq 0.4$ are shown in panel **(b,c).**

Figure 5. Genes ranked according to biological and statistical evidence

Genes were ranked (orange) based on 10 equal categories which can be divided in four categories: internal replication of genetic variant in more than two cohorts (purple; CREAM-EUR, CREAM-ASN and/or 23andMe), annotation (light blue; genetic variant harboring an exonic protein altering variant or non-protein altering variant, genetic variant residing in a 5' or 3' UTR region of a gene or transcribing an RNA structure), expression (yellow; eQTL, expression in adult human ocular tissue, expression in developing ocular tissue), biology (dark yellow; ocular phenotype in mice, ocular phenotype in humans), pathways (green; DEPICT gene-set enrichtment, DEPICT gene prioritization analysis and canonical pathway analysis of IPA). We assessed genes harboring drug targets (salmon red), but did not assign a scoring point to this category.

*Only one point can be assigned in the category 'ANNOTATION', even though it has four columns (i.e. a genetic variant is located in only 1 of these four categories).

Figure 6. Schematic representation of the human eye, retinal cell types, and functional sites of associated genes

We assessed gene expression sites and/or functional target cells in the eye for all genes using our expression data and literature and data present in the public domain. The genes appear to be distributed across virtually all cell types in the neurosensory retina, in the RPE, vascular endothelium and extracellular matrix; i.e., the route of the myopic retina-to-sclera signalling cascade.

Table 1. Results of the meta-analysis of CREAM and 23andMe for the previously-identified loci and a subset of the newly-identified loci, and replication in UK Biobank

Table 1a Replication of the HapMap II index variants for refractive error per locus in the Stage 3 meta-analysis

SNP	Chr	Position	Nearest Loci And Gene(s)	Effect Allele	Other Allele	EAF EUR	EAF ASN	Z-score	Direction	P value	HetISq	HetPVal	Sample Size (N)	P value Replication
rs12193446	6	129820038	BC035400, LAMA2	A	G	0.906	NA	-19.43	--	4.21E-84	16.8	5.72E-15	150,26	4.60E-106
rs524952	15	35005886	GOLGA8B, GJD2	A	T	0.475	0.507	-17.08	--	2.28E-65	67.2	0.015	160,15	1.60E-103
rs7744813	6	73643289	KCNQ5	A	C	0.591	0.602	-14.56	--	5.43E-48	35	0.001	160,09	1.00E-75
rs11602008	11	40149305	LRRC4C	A	T	0.822	0.749	13.98	++	2.12E-44	22.5	1.56E-10	157,50	2.90E-47
rs3138141	12	56115778	BLOC1S1-	A	C	0.214	0.147	13.8	++	2.46E-43	3.2	5.05E-07	157,53	2.30E-56
rs10500355	16	7459347	RBFOX1	A	T	0.354	0.133	-13.73	--	6.49E-43	9.1	2.93E-07	160,13	2.50E-48
rs72621438	8	60178580	SNORA51, CA8	C	G	0.642	0.609	-13.14	--	2.03E-39	38.4	0.006	160,12	1.80E-49
rs1550094	2	233385396	CHRNG, PRSS56	A	G	0.701	0.705	12.74	++	3.64E-37	26.3	0.003	159,42	4.10E-59
rs2908972	17	11407259	SHISA6	A	T	0.415	0.484	-11.13	--	9.46E-29	23	0.254	160,12	6.10E-29
rs7829127	8	40726394	ZMAT4	A	G	0.792	0.897	-10.91	--	1.02E-27	15.9	2.77E-04	160,13	3.10E-22
rs6495367	15	79375347	RASGRF1	A	G	0.408	0.399	-10.2	--	1.95E-24	0	0.667	160,14	7.20E-37
rs11145465	9	71766593	TJP2	A	C	0.212	NA	-9.55	--	1.35E-21	46.3	0.1722	153,17	1.00E-10
rs1649068	10	60304864	BICC1	A	C	0.475	0.504	-9.44	--	3.77E-21	0	0.712	160,14	7.50E-11
rs7692381	4	81903049	C4orf22, BMP3	A	G	0.763	0.63	9.4	++	5.55E-21	0	0.013	160,13	7.50E-13
rs56075542	2	146882415	BC040861,	T	G	0.552	0.472	-8.99	--	2.39E-19	13.9	0.001	159,47	1.30E-18
rs7895108	10	79061458	KCNMA1	T	G	0.351	0.118	-8.87	--	7.56E-19	32.8	0.021	160,14	1.10E-27
rs7624084	3	141093285	ZBTB38	T	C	0.568	0.633	-8.81	--	1.24E-18	18.5	0.018	160,15	6.50E-17
rs62070229	17	31227593	MYO1D, TMEM98	A	G	0.807	0.874	8.58	++	9.64E-18	0	0.416	156,57	1.30E-18
rs2855530	14	54421917	BMP4	C	G	0.507	0.474	-8.58	--	9.87E-18	41.7	0.19	160,09	4.80E-22
rs7662551	4	80537638	LOC100506035,	A	G	0.723	0.558	8.53	++	1.47E-17	19.4	0.265	160,14	6.00E-12
rs9517964	13	100717833	ZIC2,PCCA	T	C	0.589	0.786	8.42	++	3.68E-17	0	0.02	160,12	3.40E-20
rs1954761	11	105596885	GRIA4	T	C	0.371	0.377	-8.4	--	4.57E-17	0	0.911	160,12	1.20E-16
rs745480	10	85986554	LRIT2,LRIT1	C	G	0.511	0.418	8.31	++	9.26E-17	67.3	0.081	159,50	8.20E-18
rs2573081	2	178828507	PDE11A	C	G	0.524	0.538	8.21	++	2.18E-16	47.6	0.167	160,12	1.60E-29
rs17428076	2	172851936	HAT1, METAP1D	C	G	0.768	0.854	-8.18	--	2.77E-16	0	0.003	160,15	7.50E-08
rs2155413	11	84634790	DLG2	A	C	0.482	0.655	-7.76	--	8.85E-15	0	2.99E-04	159,50	1.10E-17
rs11178469	12	71275137	PTPRR	T	C	0.752	0.638	-7.4	--	1.33E-13	0	0.6989	160,13	2.60E-04
rs1858001	1	207488004	C4BPA,CD55	C	G	0.676	0.415	7.28	++	3.45E-13	59.6	0.02	160,14	6.70E-20
rs4793501	17	68718734	KCNJ2, BC039327	T	C	0.575	0.444	-7.21	--	5.53E-13	0	0.592	160,15	3.70E-12
rs7042950	9	77149837	RORB	A	G	0.732	0.392	6.8	++	1.07E-11	0	0.912	160,15	2.90E-18
rs4687586	3	53837971	CACNA1D	C	G	0.691	NA	-6.55	--	5.86E-11	0	0.605	150,21	1.60E-08
rs2753462	14	60850703	JB175233, C14orf39	C	G	0.296	0.568	-6.49	--	8.37E-11	73.9	0.05	157,35	2.00E-15
rs837323	13	101175664	PCCA	T	C	0.512	0.762	6.32	++	2.65E-10	35.6	0.213	160,14	5.30E-16
rs17382981	10	94953258	CYP26A1,MYOF	T	C	0.417	0.19	-6.31	--	2.72E-10	67.9	0.077	155,33	4.10E-07
rs79266634	16	7309047	RBFOX1	C	G	0.093	0.115	-5.93	--	3.00E-09	0	0.561	156,26	1.50E-08
rs235770	20	6761765	BMP2	T	C	0.372	0.388	-5.93	--	3.11E-09	0	0.547	157,52	4.80E-11

Table 1b Subset of the new loci harboring the smallest p-values for refractive error in the Stage 3 meta-analysis

SNP	Chr	Position	Nearest Loci And Gene(s)	Effect Allele	Other Allele	EAF EUR	EAF ASN	Z-score	Direction	P value	HetISq	HetPVal	Sample Size (N)	P value Replication
rs36024104	14	42294993	LRFN5	A	G	0.823	NA	9.09	++	9.86E-20	15.9	0.01414	152,585	2.20E-12
rs1556867	1	164213686	5S_rRNA, PBX1	T	C	0.264	0.494	-8.81	--	1.29E-18	71.1	0.06266	160,155	4.20E-17
rs2225986	1	200311910	LINC00862	A	T	0.381	0.169	-7.96	--	1.68E-15	40.2	0.196	160,152	7.50E-17
rs1207782	6	22059967	LINC00340	T	C	0.577	0.265	-7.92	--	2.47E-15	0	0.8946	160,149	4.90E-13
rs72826094	10	114801488	TCF7L2	A	T	0.799	0.838	7.88	++	3.20E-15	64.5	0.09323	156,825	4.90E-02
rs297593	2	157363743	GPD2	T	C	0.286	0.257	-7.82	--	5.45E-15	0	0.5285	159,461	7.80E-11
rs5442	12	6954864	GNB3	A	G	0.068	NA	-7.82	--	5.48E-15	8.8	0.03693	146,217	1.20E-33
rs10880855	12	46144855	ARID2	T	C	0.507	0.464	-7.78	--	7.35E-15	0	0.9683	160,144	4.80E-08
rs2150458	21	47377296	PCBP3, COL6A1	A	G	0.455	0.641	7.74	++	1.04E-14	55.7	0.1329	160,151	1.80E-13
rs12898755	15	63574641	APH1B	A	G	0.245	0.456	7.53	++	4.98E-14	7.9	0.2974	159,506	1.40E-16
rs7122817	11	117657679	DSCAML1	A	G	0.507	0.662	7.51	++	5.73E-14	73.8	0.05077	160,147	1.10E-10
rs10511652	9	18362865	SH3GL2, ADAMTSL1	A	G	0.416	0.445	7.36	++	1.91E-13	44.8	0.1782	160,149	3.50E-18
rs11101263	10	49414181	FRMPD2	T	C	0.258	0.105	-7.33	--	2.33E-13	0	0.3477	160,155	2.20E-13
rs11118367	1	219790221	LYPLAL1	T	C	0.482	0.630	-7.29	--	3.16E-13	0	0.8576	160,141	1.20E-13
rs9395623	6	50757699	TFAP2D, TFAP2B	A	T	0.315	0.381	7.25	++	4.16E-13	0	0.9579	160,151	2.20E-10
rs284816	8	53362145	ST18, FAM150A	A	G	0.163	0.198	-7.21	--	5.52E-13	0	0.9242	160,140	1.60E-08
rs12965607	18	47391025	MYO5B	T	G	0.857	0.923	7.07	++	1.52E-12	20.8	0.01674	157,604	8.10E-16
rs7747	4	80827062	ANTXR2	T	C	0.202	0.093	7.03	++	2.05E-12	5.4	0.01267	150,327	7.70E-16
rs12451582	17	54734643	NOG, C17orf67	A	G	0.369	0.308	7.02	++	2.22E-12	0	0.5925	160,155	8.80E-18
rs80253120	17	14138507	CDRT15	T	C	0.626	0.723	6.97	++	3.25E-12	58.6	0.12	156,054	7.20E-11
rs7968679	12	9313304	PZP	A	G	0.700	0.894	6.95	++	3.65E-12	0	0.01951	160,076	4.20E-10
rs11202736	10	90142203	RNLS	A	T	0.717	0.762	-6.92	--	4.53E-12	0	0.4007	160,150	9.40E-07
rs72655575	8	60556509	SNORA51, CA8	A	C	0.201	0.124	6.87	++	6.54E-12	0	0.8811	156,566	7.10E-07
rs1790165	11	131928971	NTM	A	C	0.411	0.283	6.85	++	7.17E-12	0	0.003708	160,131	1.80E-10
rs511217	11	30029948	METTL15, KCNA4	A	T	0.738	0.729	-6.79	--	1.10E-11	0	0.3626	160,143	1.40E-17

We identified 140 loci for refractive error with genome-wide significance ($P < 5 \times 10{-}8$) on the basis of the meta-analyses of the genome-wide single-variant linear regressions performed in 160,420 participants of mixed ancestries (CREAM-ASN, CREAM-EUR and 23andMe). Shown are the replication of the previously found loci from HapMap II and a subset of the new loci harboring the smallest *P* values. For each locus, represented by an index variant (the variant with smallest p-value in that locus), Effect Allele, Other Allele, effect allele frequencies per ancestry (EAF AZN and EAF EUR), effect size (Z-score), direction of the effect (Direction), the *P* value, heterogeneity *I* square (HetISq), heterogeneity P value (HetPval), Sample Size (N), and *P* value of the replication in UK Biobank are shown (Full table: Supplementary Table 2). Chr, chromosome; EAF, effect allele frequency; ASN, Asian; EUR, European; GWS, genome wide significant.

Table 2. Genetic correlation for refractive error between Europeans and East Asians

Sample 1	Sample 2	Genetic effect correlation (*pge*) [a]	Standard error *pge*	P value *pge*	Genetic impact correlation (*pgi*) [a]	Standard error *pgi*	P value *pgi*
EUR CREAM	EAS CREAM	0.804	0.041	1.83E-06	0.888	0.061	0.065
EUR 23andMe	EAS CREAM	0.788	0.041	2.48E-07	0.865	0.054	0.014

Abbreviations: EUR, European; EAS, East Asian.

[a] P-value relates to a test of the null hypothesis that *pge*=1 or *pgi*=1.

We calculated the genetic correlation of effect (pge) and impact (pgi) using Popcorn to compare the genetic associations between Europeans (CREAM-EUR, N= 44,192; 23andMe, N=104,292) and East Asians (CREAM-ASN, N= 9,826). Reference panels for Popcorn were constructed using genotype data for 503 EUR and 504 EAS individuals sequenced as part of the 1000 Genomes Project. SNPs used had a MAF of at least 5% in both populations, resulting in a final set of 3,625,602SNPs for the analyses using the 23andMe GWAS sample and 3,642,928 SNPs for those using the CREAM-EUR sample. These findings support a largely common genetic predisposition to refractive error and myopia in Europeans and Asians, although ancestry-specific risk alleles may exist.

ONLINE METHODS

Ethics Statement

All human research was approved by the relevant institutional review boards and conducted according to the Declaration of Helsinki. All CREAM participants provided written informed consent; all 23andMe applicants provided informed consent online, and answered surveys according to 23andMe's human subjects protocol, which was reviewed and approved by Ethical & Independent Review Services, an AAHRPP-accredited institutional review board. The UK Biobank received ethical approval from the National Health Service National Research Ethics Service (reference 11/NW/0382).

Study data

The study populations were participants of the Consortium for Refractive Error and Myopia (CREAM) comprising of 41,793 individuals with European ancestry from 26 cohorts (CREAM-EUR) and 11,935 individuals with Asian ancestry from 8 studies (CREAM-ASN); and customers of the 23andMe genetic testing company who gave informed consent for inclusion in research studies consisting of 104,293 individuals (2 cohorts of individuals with European ancestry, N = 12,128 and N= 92,165, respectively). All participants included in this analysis from CREAM and 23andMe were aged 25 years or older. Participants with conditions that could alter refraction, such as cataract surgery, laser refractive procedures, retinal detachment surgery, keratoconus as well as ocular or systemic syndromes were excluded from the analyses. Recruitment and ascertainment strategies varied per study (Supplementary Table 1a,b, and Supplementary Note). Refractive error represented by measurements of refraction and analyzed as spherical equivalent (SphE =spherical refractive error + 1/2 cylinder refractive error) was the outcome variable for CREAM; myopic refractive error represented by self-reported age of diagnosis of myopia (AODM) for 23andMe[27].

Genotype calling and imputation

Samples were genotyped on different platforms and study specific quality control measures of the genotyped variants were implemented before association analysis (Supplementary Table 1b). Genotypes were imputed using the appropriate ancestry-matched reference panel for all cohorts from the 1000 Genomes Project (Phase I version 3, March 2012 release) with either minimac[100] or IMPUTE[101]. The metrics for pre-imputation quality control varied amongst studies, but genotype call rate thresholds were set at high level (≥0.95 for both CREAM and 23andMe). These metrics were similar to our previous GWAS analyses[26,27]; details per cohort can be found in Supplementary Table 1b.

GWAS per study

For each CREAM cohort, a single marker analysis for the SphE (in diopters) phenotype was carried out using linear regression adjusting for age, sex and up to the first five principal components. All non-family-based cohorts removed one of each pair of relatives (after detection using either GCTA or IBS/IBD analysis). In family-based cohorts, a score test-based association was used to adjust for within-family relatedness[102]. For the 23andMe participants, Cox proportional hazards analysis testing AODM as the dependent variable were performed as previously described[27], with *P* calculated using a likelihood ratio test for the single marker genotype term. We used an additive SNP allelic effect model for all analyses.

Centralized quality control per study

After individual GWAS, all studies underwent a second round of quality control (QC). Quantile-quantile, effect allele frequency, *P* – *Z* test, standard error – sample size, and genomic control inflation factor plots were generated for each individual cohort using EasyQC[103] (Supplementary Figure 2.1 (Supplementary Figure 2.1.1 and 2.1.2), 2.2 (Supplementary Figure 2.2.1 – 2.2.5), 2.3 (Supplementary Figure 2.3.1 and 2.3.2). All analytical issues discovered during this QC step were resolved per individual cohort.

GWAS meta-analyses

The GWAS meta-analyses were performed in three stages (Supplementary Figure 1). In *Stage 1*, European (CREAM-EUR, N=44,192) and Asian (CREAM-ASN, N=11,935) participants from the CREAM cohort were meta-analysed separately. Subsequently, all CREAM cohorts (CREAM-ALL) were meta-analysed. Variants with MAF < 1% or imputation quality score < 0.3 (info metric of IMPUTE) or Rsq < 0.3 (minimac) were excluded. A fixed effects inverse variance-weighted meta-analysis was performed using METAL[104]. 1,063 variants clustering in 24 loci (Supplementary Table 2) were genome-wide significant (P=5.0x10^{-8}). All 37 loci that were previously found by CREAM and 23andMe using genotype data imputed to the HapMap II reference panel were replicated (*pBonferroni* 1.85x10^{-3}), and 36 of the 37 were genome-wide significant (Supplementary Table 2)[26,27]. In *Stage 2*, a meta-analysis of the two 23andMe cohorts ($N_{23andMe\_V2}$=12,128; $N_{23andMe\_V3}$=92,165) was performed, using similar filtering but a lower MAF threshold (< 0.5%). A total of 5,205 genome-wide significant variants clustered in 112 loci (Supplementary Table 2).

In *Stage 3*, CREAM-ALL and 23andMe samples were combined using a fixed effects meta-analysis based on *P* value and direction of effect. In all stages, each genetic variant had to be represented by at least half of the entire study population and at least represented by 13 cohorts in CREAM and one cohort in 23andMe. For SNPs with high heterogeneity (at $P < 0.05$), we also performed a random effects meta-analysis using METASOFT[50]. We chose a different weighting scheme due to the differences in effect size scaling; 23andMe used a less accurate phenotype variable (AODM); i.e. the *effective* sample size of the 23andMe was approximately equivalent to the *effective* sample size of CREAM-ALL (Figure 2b), thus weighting by ($1/\sqrt{n_{effective}}$) yielded a final weighting ratio of 1:1[105]. Genome-wide statistical significance was defined at $P < 5.0 \times 10^{-8}$ [106].

All three meta-analysis stages were performed under genomic control. Study specific and meta-analysis lambda (λ) estimates are shown in Supplementary Figure 6; to check for confounding biases (e.g. cryptic relatedness and population stratification), LD score intercepts from LD score regressions per ancestry were constructed (Supplementary Figure 7)[30]. To check the robustness of signals, we performed a conventional random effects models using METASOFT, fixed effects models

weighted on sample size and on weights estimated from standard error per allele tested using METAL (Supplementary Table 2 and Supplementary Table 3).

Manhattan (modified version of package 'qqman'), regional, box, and forest plots were made using R version 3.2.3 and LocusZoom[107]. An overview of the Hardy Weinberg *P* of all index variants per cohort can be found in Supplementary Table 4. The comparison between refractive error and age-of-onset was performed using the LDSC program[30].

Population stratification and heritability calculations

Each study assessed the degree of genetic admixture and stratification in their study participants through the use of principal components. Homogeneity of participants was assured by removal of all individuals whose ancestry did not match the prevailing ancestral group. We used genomic inflation factors to control for admixture and stratification, and performed genomic-controlled meta-analysis to account for the effects of any residual heterogeneity. To further distinguish between inflation from a true polygenic signal and population stratification, we examined the relationship between test statistics and linkage disequilibrium (LD) with LDSC. CREAM-EUR, CREAM-ASN and 23andMe were evaluated separately; variants not present in HapMap3 and MAF < 1% were excluded. SNP heritability estimates were calculated using LDSC for the same set of genetic variants.

Locus definition and annotation

All study effect size estimates were oriented to the positive strand of the NCBI Build 37 reference sequence of the human genome. The index variant of a locus was defined as the variant with the lowest *P* in a region spanning a 100 kb window of the most outer genome-wide significant variant of that same region. We annotated all index variants using the web-based version of ANNOVAR[108] based on UCSC Known Gene Database[35]. For variants within the coding sequence or 5' or 3' untranslated regions of a gene, that gene was assigned to the index variant (note that this led to more than 1 gene being assigned to variants located within the transcription units of multiple, overlapping genes). For variants in intergenic regions, the nearest 5' gene and the nearest 3' gene were assigned to the variant.

Index variants were annotated to functional RNA elements when described as such in the UCSC Known Gene Database. We used conservation (PhyloP[109]) and prediction tools (SIFT[39], Mutation Taster[110], align GVGD[40,41], PolyPhen-2[36]) to predict the pathogenicity of protein-altering exonic variants.

Conditional signal analysis

We performed conditional analysis to identify additional independent signals nearby the index variant at each locus, using GCTA-COJO[32]. We transformed the Z-scores of the summary statistics to beta's using the following formula: Standard Error = $\sqrt{1/2 * N * MAF(1 - MAF)}$. We performed the GCTA-COJO analysis[32], utilizing summary-level statistics from the meta-analysis on all cohorts. Linkage disequilibrium (LD) between variants was estimated from the Rotterdam Study I-III.

Replication in UK Biobank

The UK Biobank Eye & Vision (UKEV) Consortium performed a GWAS of refractive error in 95,505 participants of European ancestry aged 37-73 year with no history of eye disorders[33]. Refractive error was measured using an autorefractor; SphE was calculated per eye and averaged between the two eyes. To account for relatedness a mixed model analysis with BOLT-LMM was used[111], including age, gender, genotyping array, and the first 10 principal components as covariates. Analysis was restricted to markers present on the HRC reference panel[112]. We performed lookups for all independent genetic variants identified in our Stage 3 meta-analysis and conditional analysis. For 16 variants not present in UKEV, we performed lookups for a surrogate variant in high LD (r^2 >0.8). When more than one potential surrogate variant was available, the variant in strongest LD with the index variant was selected. Six variants were not available for replication: one variant (rs188159083) was not present on the array nor was a surrogate available in UKEV and five variants showed evidence of departure from HWE (HWE exact test $P<3.0x10^{-4}$).

Post-GWAS analyses

We performed two gene-based tests to identify additional significant genes not found in the single variant analysis. First, we applied the gene-based test implemented in fastBAT[42] to the per-variant summary statistics of the meta-analysis of all European cohorts (23andMe and CREAM-EUR). We used the default parameters (all variants in or within 50kb of a gene) and focused on variants with a gene-based $P < 2 \times 10^{-6}$ (Bonferroni correction based on 25,000 genes) and the per-variant $P > 5 \times 10^{-8}$. Secondly, we applied another gene-based test in EUGENE[43] which only includes variants which are eQTLs (GTex, blood[113]). EUGENE tests an hypothesis predicated on eQTLs as key drivers of the association signal. eQTLs within 50kb of a gene were included in the test. Genes with EUGENE $P < 2 \times 10^{-6}$ (and not found in the single variant analysis) were considered to be significant. Finally, we used functional annotation information from genome-wide significant loci to reweigh results using fgwas (version 0.3.64[44]). Fgwas incorporates functional annotation (e.g. DNase I-hypersensitive sites in various tissues and 3'UTR regions) to reweight data from GWAS, and uses a Bayesian model to calculate a posterior probability of association. This approach is able to identify risk loci that otherwise might not reach the genome-wide significance threshold in standard GWAS. Details about this approach can be found in Supplementary Note.

Refractive errors and myopia risk prediction

To assess the risk of the entire range of refractive errors, we computed polygenic risk scores (PGRS) for the population-based Rotterdam Studies (RS) I, RS-II and RS-III using the *P* and *Z* scores from a meta-analysis on CREAM-ALL and 23andMe, excluding the RS I-III cohorts. Only variants with high imputation quality (IMPUTE info score > 0.5 or minimac Rsq > 0.8) and MAF > 1% were considered. *P*-based clumping was performed with PLINK[114], using an r^2 threshold of 0.2 and a physical distance threshold of 500 kb, excluding the MHC region. This resulted in a total of 243,938 variants. For each individual in RS-I, RS-II and RS-III (N = 10,792), PGRS were calculated using the --score command in PLINK across strata of *P* thresholds: 5.0×10^{-8}, 5.0×10^{-7}, 5.0×10^{-6}, 5.0×10^{-5}, 5.0×10^{-4}, 0.005, 0.01, 0.05, 0.1, 0.5, 0.8 and 1.0. The proportion of variance explained by each PGRS model was

calculated as the difference in the R^2 between two regression models; one where SphE was regressed on age, sex, the first five principal components, and the other also including the PGRS as an additional covariate. Subsequently, AUCs were calculated for myopia (SphE ≤ -3 SD) versus hyperopia (SphE ≥ +3 SD).

Genetic correlation between ancestries

We used Popcorn[48] to investigate ancestry-related differences in the genetic architecture of refractive error and myopia. Popcorn takes summary GWAS statistics from two populations and LD information from ancestry-matched reference panels, and computes genetic correlations by implementing a weighted likelihood function that accounts for the inflation of Z scores due to LD. Pairwise analyses were carried out using the GWAS summary statistics from 23andMe (*N* = 104,292), CREAM-EUR (*N* = 44,192) and CREAM-EAS (*N* = 9,826) meta-analyses. Only SNPs with MAF ≥ 5% were included, resulting in a final set of 3,625,602 SNPs for analyses involving 23andMe and 3,642,928 SNPs for the CREAM-EUR versus CREAM-EAS analysis. Reference panels were constructed using genotype data from 503 European and 504 East Asian individuals sequenced as part of the 1000 Genomes Project (release 2013-05-02 downloaded from: ftp.1000genomes.ebi.ac.uk). The reference panel VCF files were filtered using PLINK[114] to remove indels, strand-ambiguous variants, variants without an "rs" id prefix, and variants located in the MHC region on chromosome 6 (chr6:25,000,000-33,500,000; Build 37).

Analysis between phenotypes

To evaluate consistency of genotypic effects across studies that employed different phenotype definitions, we compared effect sizes from GWAS studies of either SphE or AODM in Europeans, i.e. CREAM-EUR (*N* = 44,192) or 23andMe (*N* = 104,293) respectively. Marker-wise additive genetic effect sizes (in units diopters per copy of the risk allele) for SphE were compared against those (in units log(HR) per copy of the risk allele) for AODM. Data was visualised using R. Genetic correlation

between the two phenotypes SphE and AODM was calculated using LD score regression. This analysis included all common SNPs (MAF > 0.01) present in HapMap3.

Evidence for functional involvement

In order to rank genes according to biological plausibility, we scored annotated genes based on our own findings and published reports for a potential functional role in refractive error. Points were assigned for each gene on the basis of 10 categories (details on the methodology per category are provided in Supplementary Note): internal replication of index genetic variants in the individual cohort GWAS analyses through Bonferroni corrections (CREAM-ASN, CREAM-EUR and 23andMe; *pBonferroni* 1.19×10^{-4}), evidence for eQTL using the FUMA[32] and extensive look-ups in GtEx, evidence of expression in the eye in developmental and adult ocular tissues, presence of an eye phenotype in knock-out mice (MGI and IMPC database), presence of an eye phenotype in humans (OMIM; see URLs, DisGeNET[115]), location in a functional region of a gene (wANNOVAR; see URLs), presence of the gene in a significant enriched functional pathway with false discovery rate < 0.05 (DEPICT[49]), presence of the gene in the gene priority analysis of DEPICT with false discovery rate < 0.05 and the presence of the gene in the canonical pathway analysis of Ingenuity Pathway Analysis (IPA; See URLs). Furthermore, we performed a systematic search for each gene to assess its potential as a drug target (SuperTarget[116], STITCH[117], DrugBank[118], PharmaGkb[119]). All information derived from this study and literature were used to annotate genes to retinal cell types.

Genetic pleiotropy

To investigate overlap of genes with other common traits, we performed a look-up in the GWAS catalog using FUMA. Multiple testing correction (i.e. Benjamini-Hochberg) was performed. Traits were significantly associated when adjusted $P \leq 0.05$ and the number of genes that overlap with the GWAS catalog gene sets was ≥ 2.

Data availability statement

The summary statistics of the Stage 3 meta-analysis are included in Supplementary Data 3 of this published article. In order to protect the privacy of the participants in our cohorts, further summary statistics of Stage 1 (CREAM) and Stage 2 (23andMe) will be available upon request. Please contact c.c.w.klaver@erasmusmc.nl (CREAM) and/or apply.research@23andMe.com (23andMe) for more information and to access the data.

Methods only references

100. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-9 (2012).
101. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-13 (2007).
102. Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans. *Am J Hum Genet* **81**, 913-26 (2007).
103. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nat Protoc* **9**, 1192-212 (2014).
104. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
105. Zaykin, D.V. Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *J Evol Biol* **24**, 1836-41 (2011).
106. Dudbridge, F. & Gusnanto, A. Estimation of significance thresholds for genomewide association scans. *Genet Epidemiol* **32**, 227-34 (2008).
107. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-7 (2010).
108. Yang, H. & Wang, K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. *Nat Protoc* **10**, 1556-66 (2015).
109. Cooper, G.M. *et al.* Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res* **15**, 901-13 (2005).
110. Schwarz, J.M., Rodelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods* **7**, 575-6 (2010).
111. Loh, P.R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* **47**, 284-90 (2015).
112. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* **48**, 1279-83 (2016).
113. Consortium, G.T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
114. Chang, C.C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
115. Bauer-Mehren, A., Rautschka, M., Sanz, F. & Furlong, L.I. DisGeNET: a Cytoscape plugin to visualize, integrate, search and analyze gene-disease networks. *Bioinformatics* **26**, 2924-6 (2010).
116. Gunther, S. *et al.* SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic Acids Res* **36**, D919-22 (2008).
117. Kuhn, M. *et al.* STITCH 4: integration of protein-chemical interactions with user data. *Nucleic Acids Res* **42**, D401-7 (2014).
118. Wishart, D.S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res* **34**, D668-72 (2006).
119. Whirl-Carrillo, M. *et al.* Pharmacogenomics knowledge for personalized medicine. *Clin Pharmacol Ther* **92**, 414-7 (2012).

The CREAM Consortium

Tin Aung[1,2], Joan E. Bailey-Wilson[3], Paul Nigel Baird[4], Amutha Barathi Veluchamy[1,5], Ginevra Biino[6], Kathryn P. Burdon[7], Harry Campbell[8], Li Jia Chen[9], Peng Chen[2], Wei Chen[10], Ching-Yu Cheng[11,12], Emily Chew[13], Jamie E. Craig[7], Phillippa M. Cumberland[14], Margaret M. Deangelis[15], Cécile Delcourt[16], Xiaohu Ding[17], Angela Döring[18], Cornelia M. van Duijn[19], David M. Evans[20,21], Qiao Fan[12], Lindsay Farrer[15], Sheng Feng[22], Brian Fleck[23], Rhys D. Fogarty[7], Jeremy R. Fondran[24], Maurizio Fossarello[25], Paul J. Foster[26], Puya Gharahkhani[27], Christian Gieger[18], Adriana I. Iglesias[19,28,29], Jeremy A. Guggenheim[30], Xiaobo Guo[17,31], Toomas Haller[32], Christopher J. Hammond[33], Caroline Hayward[34], Mingguang He[4,17], Alex W. Hewitt[4,35,36], René Höhn[37,38], S. Mohsen Hosseini[39], Laura D. Howe[21,40], Pirro G. Hysi[33], Robert P. Igo Jr.[24], Sudha K. Iyengar[13,24,41], Sarayut Janmahasatian[24], Vishal Jhanji[9], Jost B. Jonas[42,43], Mika Kähönen[44], Jaakko Kaprio[45,46], John P. Kemp[21], Kay-Tee Khaw[47], Anthony P. Khawaja[26,47], Chiea-Chuen Khor[2,15,48,49], Caroline C. W. Klaver[19,29,50], Barbara E. Klein[51], Ronald Klein[51], Eva Krapohl[52], Jean-François Korobelnik[53,54], Jonathan H. Lass[24,41], Kris Lee[51], Elisabeth M. van Leeuwen[19,29], Terho Lehtimäki[55,56], Shi-Ming Li[43], Yi Lu[27], Robert N. Luben[47], Stuart MacGregor[27], David A. Mackey[4,35,36], Kari-Matti Mäkelä[55], Nicholas G. Martin[57], George McMahon[21], Akira Meguro[58], Thomas Meitinger[59,60], Andres Metspalu[32], Evelin Mihailov[32], Paul Mitchell[61], Masahiro Miyake[62], Nobuhisa Mizuki[58], Margaux Morrison[15], Vinay Nangia[63], Songhomitra Panda-Jonas[63], Chi Pui Pang[9], Olavi Pärssinen[64,65], Andrew D. Paterson[39], Norbert Pfeiffer[38], Mario Pirastu[66], Robert Plomin[52], Ozren Polasek[8,67], Jugnoo S. Rahi[14,26,68], Olli Raitakari[69,70], Taina Rantanen[65], Janina S. Ried[18], Igor Rudan[8], Seang-Mei Saw[48,71], Maria Schache[4], Ilkka Seppälä[55], Rupal L. Shah[30], George Davey Smith[21], Dwight Stambolian[72], Beate St Pourcain[21,73], Claire L. Simpson[3,74], E-Shyong Tai[48], Pancy O. Tam[9], Milly S. Tedja[19,29], Yik-Ying Teo[48,75], J. Willem L. Tideman[19,29], Nicholas J. Timpson[21], Simona Vaccargiu[66], Zoran Vatavuk[76], Virginie J.M. Verhoeven[19,28,29], Veronique Vitart[34], Jie Jin Wang[4,61], Ningli Wang[43], Nick J. Wareham[77], Juho Wedenoja[45,78], Cathy Williams[79], Katie M. Williams[33], James F. Wilson[8,34], Robert Wojciechowski[3,80,81], Ya Xing Wang[43], Tien-Yin Wong[82,83], Alan F. Wright[34], Jing Xie[4], Liang Xu[43],

Kenji Yamashiro[62], Maurice K.H. Yap[84], Seyhan Yazar[36], Shea Ping Yip[85], Nagahisa Yoshimura[62],

Alvin L. Young[9], Terri L. Young[51], Jing Hua Zhao[77], Wanting Zhao[12,86], Xiangtian Zhou[10]

Affiliations

1. Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
2. Department of Ophthalmology, National University Health Systems, National University of Singapore, Singapore.
3. Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA.
4. Centre for Eye Research Australia, Ophthalmology, Department of Surgery, University of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Australia.
5. Duke-NUS Medical School, Singapore, Singapore.
6. Institute of Molecular Genetics, National Research Council of Italy, Sassari, Italy.
7. Department of Ophthalmology, Flinders University, Adelaide, Australia.
8. Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK.
9. Department of Ophthalmology and Visual Sciences, The Chinese University of Hong Kong, Hong Kong Eye Hospital, Kowloon, Hong Kong.
10. School of Ophthalmology and Optometry, Eye Hospital, Wenzhou Medical University, China.
11. Ocular Epidemiology Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
12. Centre for Quantitative Medicine, DUKE-National University of Singapore, Singapore.
13. Department of Genetics, Case Western Reserve University, Cleveland, Ohio, USA.
14. Great Ormond Street Institute of Child Health, University College London, London, UK.
15. Department of Ophthalmology and Visual Sciences, John Moran Eye Center, University of Utah, Salt Lake City, Utah, USA.
16. Université de Bordeaux, Inserm, Bordeaux Population Health Research Center, team LEHA, UMR 1219, F-33000 Bordeaux, France.
17. State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China.
18. Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany.
19. Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands.
20. Translational Research Institute, University of Queensland Diamantina Institute, Brisbane, Queensland, Australia.
21. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.
22. Department of Pediatric Ophthalmology, Duke Eye Center For Human Genetics, Durham, North Carolina, USA.
23. Princess Alexandra Eye Pavilion, Edinburgh, UK.
24. Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, USA.
25. University Hospital 'San Giovanni di Dio', Cagliari, Italy.
26. NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK.
27. Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
28. Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands.
29. Department of Ophthalmology, Erasmus Medical Center, Rotterdam, The Netherlands.
30. School of Optometry & Vision Sciences, Cardiff University, Cardiff, UK.
31. Department of Statistical Science, School of Mathematics, Sun Yat-Sen University, Guangzhou, China.
32. Estonian Genome Center, University of Tartu, Tartu, Estonia.
33. Section of Academic Ophthalmology, School of Life Course Sciences, King's College London, London, UK.

34. MRC Human Genetics Unit, MRC Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh, UK.
35. Department of Ophthalmology, Menzies Institute of Medical Research, University of Tasmania, Hobart, Australia.
36. Centre for Ophthalmology and Visual Science, Lions Eye Institute, University of Western Australia, Perth, Australia.
37. Department of Ophthalmology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland.
38. Department of Ophthalmology, University Medical Center Mainz, Mainz, Germany.
39. Program in Genetics and Genome Biology, Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada.
40. School of Social and Community Medicine, University of Bristol, Bristol, UK.
41. Department of Ophthalmology and Visual Sciences, Case Western Reserve University and University Hospitals Eye Institute, Cleveland, Ohio, USA.
42. Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University of Heidelberg, Mannheim, Germany.
43. Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China.
44. Department of Clinical Physiology, Tampere University Hospital and School of Medicine, University of Tampere, Tampere, Finland.
45. Department of Public Health, University of Helsinki, Helsinki, Finland.
46. Institute for Molecular Medicine Finland FIMM, HiLIFE Unit, University of Helsinki, Helsinki, Finland.
47. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
48. Saw Swee Hock School of Public Health, National University Health Systems, National University of Singapore, Singapore.
49. Division of Human Genetics, Genome Institute of Singapore, Singapore.
50. Department of Ophthalmology, Radboud University Medical Center, Nijmegen, The Netherlands.
51. Department of Ophthalmology and Visual Sciences, University of Wisconsin–Madison, Madison, Wisconsin, USA.
52. MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK.
53. Université de Bordeaux, Bordeaux, France.
54. INSERM (Institut National de la Santé Et de la Recherche Médicale), ISPED (Institut de Santé Publique d'Épidémiologie et de Développement), Centre INSERM U897-Epidemiologie-Biostatistique, Bordeaux, France.
55. Department of Clinical Chemistry, Finnish Cardiovascular Research Center-Tampere, Faculty of Medicine and Life Sciences, University of Tampere.
56. Department of Clinical Chemistry, Fimlab Laboratories, University of Tampere, Tampere, Finland.
57. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
58. Department of Ophthalmology, Yokohama City University School of Medicine, Yokohama, Kanagawa, Japan.
59. Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany.
60. Institute of Human Genetics, Klinikum rechts der Isar, Technische Universität München, Munich, Germany.
61. Department of Ophthalmology, Centre for Vision Research, Westmead Institute for Medical Research, University of Sydney, Sydney, Australia.
62. Department of Ophthalmology and Visual Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan.
63. Suraj Eye Institute, Nagpur, Maharashtra, India.
64. Department of Ophthalmology, Central Hospital of Central Finland, Jyväskylä, Finland.

65. Gerontology Research Center, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland.
66. Institute of Genetic and Biomedic Research, National Research Council, Cagliari, Italy.
67. Faculty of Medicine, University of Split, Split, Croatia.
68. Ulverscroft Vision Research Group, University College London, London, UK.
69. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland.
70. Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland.
71. Myopia Research Group, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.
72. Department of Ophthalmology, University of Pennsylvania, Philadelphia, Pennsylvania, USA.
73. Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.
74. Department of Genetics, Genomics and Informatics, University of Tennessee Health Sciences Center, Memphis, Tenessee.
75. Department of Statistics and Applied Probability, National University of Singapore, Singapore.
76. Department of Ophthalmology, Sisters of Mercy University Hospital, Zagreb, Croatia.
77. MRC Epidemiology Unit, Institute of Metabolic Sciences, University of Cambridge, Cambridge, UK.
78. Department of Ophthalmology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland.
79. Department of Population Health Sciences, Bristol Medical School, Bristol, UK.
80. Department of Epidemiology and Medicine, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA.
81. Wilmer Eye Institute, Johns Hopkins Medical Institutions, Baltimore, Maryland, USA.
82. Academic Medicine Research Institute, Singapore.
83. Retino Center, Singapore National Eye Centre, Singapore, Singapore.
84. Centre for Myopia Research, School of Optometry, The Hong Kong Polytechnic University, Hong Kong, Hong Kong.
85. Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, Hong Kong.
86. Statistics Support Platform, Singapore Eye Research Institute, Singapore National Eye Centre, Singapore.

SUPPLEMENTS

Tedja M.S. et al. Large genome-wide meta-analysis highlights light-induced signaling as a driver for refractive error.

Supplementary Note:

1: Study Populations and Acknowledgements

2: Methods per subfields of prioritization

3: Dopamine pathway look-ups

4: Phenotyping in CREAM and 23andMe

5: Gene-based test - fgwas

Supplementary Figures:

Supplementary Figure 1: Study design flow-chart

Supplementary Figure 2.1: EasyQC plots per cohort – CREAM-ASN

Supplementary Figure 2.1.1: EasyQC QQ, P-Z, AF-plots per cohort in CREAM-ASN

Supplementary Figure 2.1.2: EasyQC λ-N and SE-N plots per cohort in CREAM-ASN

Supplementary Figure 2.2: EasyQC plots per cohort – CREAM-EUR

Supplementary Figure 2.2.1: EasyQC QQ-plots per cohort in CREAM-EUR

Supplementary Figure 2.2.2: EasyQC PZ-plots per cohort in CREAM-EUR

Supplementary Figure 2.2.3: EasyQC AF-plots per cohort in CREAM-EUR

Supplementary Figure 2.2.4: EasyQC λ-N plots per cohort in CREAM-EUR

Supplementary Figure 2.2.5: EasyQC SE-N plots per cohort in CREAM-EUR

Supplementary Figure 2.3: EasyQC plots per cohort – 23andMe

Supplementary Figure 2.3.1: EasyQC QQ, P-Z, AF-plots per cohort in 23andMe

Supplementary Figure 2.3.2: EasyQC λ-N and SE-N plots per cohort in 23andMe

Supplementary Figure 3: Boxplot of effect sizes per cohort

Supplementary Figure 4: Imputation Quality of genetic variants of Stage 3

Supplementary Figure 5: Manhattan plots and QQ plots per stage (1-3)

Supplementary Figure 6: Lambdas of all cohorts and meta-analyses

Supplementary Figure 7: LD-score regressions per ancestry of CREAM and 23andMe

Supplementary Figure 8: Effects comparison SE and AODM using different p-value thresholds

Supplementary Figure 9: Top canonical pathways of Ingenuity Pathway Analysis

Supplementary Figure 10: Locus Zoom plots of RNA gene regions

Supplementary Figure 11: Locus Zoom plots of dopamine pathway regions

Supplementary Figure 12: GWAS catalog comparison of refractive error genes with other diseases

Supplementary Tables:

Supplementary Table 1a: Descriptives per cohort

Supplementary Table 1b: Phenotyping and imputation methods per cohort

Supplementary Table 2: Stage 1 – 3 Meta-Analyses and Conditional Analysis (separate file)

Supplementary Table 3: Index Variants of FE and RE meta-analyses (separate file)

Supplementary Table 4: HWE P value per cohort TopSNPs Stage 3 (separate file)

Supplementary Table 5: Index SNPs HapMap II from CREAM and 23andMe

Supplementary Table 6: LD score regression analysis and heritability explained by common SNPs

Supplementary Table 7: Exonic genetic variants - protein altering

Supplementary Table 8: Post-GWAS analyses - additional genes

Supplementary Table 9: Predictive power of the polygenic score in Rotterdam Studies (I, II, III)

Supplementary Table 10: DEPICT analysis - gene-set enrichment (separate file)

Supplementary Table 11: Ranking of associated genes according to biological plausibility (separate file)

Supplementary Table 12: RNA genes – region look-ups

Supplementary Table 13: Genes with Human Ocular Phenotypes (separate file)

Supplementary Table 14: Genes with Mice Ocular Phenotypes (separate file)

Supplementary Table 15: Expression of Candidate Genes in Ocular Tissues (separate file)

Supplementary Table 16: Genetic variants harboring known drug target(s) (separate file)

Supplementary Table 17: Genes in Dopamine pathway – region look-ups

Supplementary Data

Supplementary Data 1: Locus Zoom Plots (separate file)

Supplementary Data 2: Forest Plots Stage 3 and Conditional Loci (separate file)

Supplementary Data 3: Summary Statistics Stage 3 Meta-analysis (separate file; SNPs $P < 0.005$)

Supplementary Note

1: Study Populations and Acknowledgements

CREAM cohort

CREAM (Consortium for Refractive Error and Myopia) was established in 2011 as a collaboration between studies with data on refractive error which had performed genome-wide association analysis based on SNP arrays. Details of each study cohort and their group-specific acknowledgements are provided below.

1958 British Birth Cohort

The 1958 British Birth Cohort[1] is a prospective population-based cohort study that initially included 17,000 newborn children whose birth was within the first week of March 1958. All participants gave informed written consent to participate in genetic association studies, and the study was approved by the South East Multi Centre Research Ethics Committee (MREC) and the Oversight Committee for the biomedical examination of the British 1958 British birth cohort. Biomedical examination protocols were approved by the South East MREC.

1958 British Birth Cohort acknowledgements: Phenotyping was funded by the Medical Research Council's Health of the Public grant (PIs Power and Strachan); the genetic studies by the Wellcome Trust (083478 to J.S.R.); some of the analysis by the National Institute for Health Research as Specialist Biomedical Research Centres in Paediatrics and Ophthalmology, partnering respectively with Great Ormond Street and Moorfields Hospitals; with additional personal funding (P.M.C) by the Ulverscroft Vision Research Group.

ALIENOR

The Alienor study is a population-based study in residents of Bordeaux, France[2]. The 963 participants, aged 73 years or more, were recruited from an ongoing population-based study (3C Study)[3]. They underwent an ophthalmological examination, including a recording of ophthalmological history, measures of visual acuity, refraction, two 45° non mydriatic colour retinal photographs (one centred on the macula, the other centred on the optic disc), measures of intraocular pressure and central corneal thickness and break-up time test. Refraction was measured first using autorefractometer (Speedy K, Luneau, France) and secondly by measuring subjective measurement, which was used in the analysis. This research followed the tenets of the Declaration of Helsinki. Participants gave written consent for the participation

in the study. The design of this study has been approved by the Ethical Committee of Bordeaux (Comité de Protection des Personnes Sud-Ouest et Outre-Mer III) in May 2006. After exclusion of subjects operated for cataract and other eye procedures and diseases that could alter refraction, 618 subjects were available, among which 529 were genotyped at the French national centre for genotyping (CNG) using Illumina Human 610-Quad BeadChip. Among them, 509 individuals had good genotype QC (individuals of European ancestry, unrelated with other individuals, without discrepancy between clinical and genetic gender and with missingness < 5%) and had imputation data. In addition, 2 subjects had missing education data, leaving 507 subjects in the statistical analysis. Imputation was performed in two steps: prephasing with SHAPEIT2, followed by imputation with IMPUTE2 using 1000 Genomes(March 2012, MACGT1) as reference panel. SNPs were used in the imputation process if call rate > 98%, HWE p-value > 1 x 10-6 , MAF> 1%. Analysis was performed using Quicktest, with adjustment on age, gender, education, PC1 and PC2 and modelling of interaction between SNP and education, using robust variance estimates. No SNP exclusion was applied on imputed SNPs.

ALIENOR acknowledgements: The Alienor study is supported by laboratoires Théa (Clermont-Ferrand, France). The Three-City study is conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the University of Bordeaux and Sanofi-Aventis. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The Three-City study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Caisse Nationale de Solidarité pour l'Autonomie, Direction Générale de la Santé, MGEN, Institut de la Longévité, Conseils Régionaux d'Aquitaine et Bourgogne, Fondation de France, Ministry of Research-INSERM Programme "Cohortes et collections de données biologiques", Agence Nationale de la Recherche ANR PNRA 2006 and LongVie 2007, and the "Fondation Plan Alzheimer" (FCS 2009-2012). Laboratoires Théa participated in the design of the Alienor study, but none of the sponsors participated in the collection, management, statistical analysis and interpretation of the data, not in the preparation, review or approval of the present manuscript.

ALSPAC (Avon Longitudinal Study of Parents and Children)

The research adhered to the tenets of the Declaration of Helsinki. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Pregnant women with an expected date of delivery between 1st April 1991 and 31st December 1992, resident in the former Avon health authority area in Southwest England, were eligible to participate in this birth cohort study. 14,541 women were recruited. Data collection has been via various methods including self-completion questionnaires sent to the mother, to her partner and after age 5 to the child; direct

assessments and interviews in a research clinic. As well as investigating the health and well-being of the of the children in the birth cohort, the health of the mothers is also an important area of investigation[4,5]. DNA has been extracted from blood samples collected as part of routine antenatal care, during attendance at ALSPAC research clinics, or from immortalized lymphoblastoid cell lines, for a total of 10,321 of the mothers. Please note that the study website contains details of all the data that is available through a fully searchable data dictionary: http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/.

ALSPAC acknowledgements: Core support for ALSPAC was provided by the UK Medical Research Council and Wellcome Trust (Grant 102215/2/13/2) and the University of Bristol; this research specifically was funded by Wellcome Trust ISSF Populations Pilot Award (508353/509506); C.W. is supported by an NIHR Fellowship (CDF-2009-02-35). We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

AREDS

The Age-Related Eye Disease Study (AREDS) was initially designed as a long-term multicenter, prospective study of the clinical course of age-related macular degeneration (AMD) and age-related cataract[6,7]. In addition to collecting natural history data, AREDS included a randomized clinical trial of high-dose vitamin and mineral supplements for AMD and a clinical trial of high-dose vitamin supplements for cataract[6-8]. Prior to study initiation, the protocol was approved by an independent data and safety monitoring committee and by the institutional review board for each clinical center. Written informed consent was obtained from all participants before enrollment in accordance with the Declaration of Helsinki. AREDS participants were 55 to 80 years of age at enrollment and had to be free of any illness or condition that would make long-term follow-up or compliance with study medications unlikely or difficult. On the basis of fundus photographs graded by a central reading center, best-corrected visual acuity and ophthalmologic evaluations, 4,757 participants were enrolled in one of several AMD categories, including persons with no AMD (control group). Visual acuity measurement of all participants was performed with the standard procedure developed for the Early Treatment of Diabetic Retinopathy Study (ETDRS). A refraction measurement was performed for participants at the randomization visit and each annual visit. For those who experience a decrease of 10 letters from baseline visual acuity, refractions were also conducted at the non-annual visits. Blood samples were collected at baseline and longitudinally as participants were send, and cell lines were established. DNA was extracted from cell

lines according to standard protocols. For the current analysis, 816 participants aged 60 and older were included from the AREDS 1a-1b population and 1506 from the AREDS 1c population. Refractive error measured by a refraction protocol at baseline enrollment into the AREDS study[6-9] was analyzed, taking the mean measured spherical equivalent (SE) across both eyes (or SE in a single eye when both eyes were not measured) as the trait of interest. Age, gender and the first two principal components (to adjust for significant population stratification) were also included as covariates.

Acknowledgements AREDS: **AREDS1a1b** and **FECD** were supported by the National Eye Institute (grants R01EY16482, R21EY015145, and P30EY11373) and by Research to Prevent Blindness and the Ohio Lions Eye Research Foundation. The investigators gratefully acknowledge the role of the clinical coordinators and investigators who collected data on FECD cases and controls. Individual investigators and sites are listed in the first publication of the FECD study[10]. Data for the AREDS1a and 1b studies was downloaded from dbGaP for analysis under a National Eye Institute data use agreement.

AREDS1c was supported by contracts from National Eye Institute/National Institutes of Health, Bethesda, MD, with additional support from Bausch & Lomb Inc, Rochester, NY. The genotyping costs were supported by the National Eye Institute (R01EY020483 to D.S.) and some of the analyses were supported by the Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health, USA. AREDS acknowledges Frederick Ferris, National Eye Institute, National Institutes of Health, Bethesda, MD; and the Center for Inherited Disease Research, Baltimore, MD where SNP genotyping was carried out. The investigators gratefully acknowledge the advice and guidance of Hemin Chin of the National Eye Institute.

Blue Mountains Eye Study (BMES)

The Blue Mountains Eye Study (BMES) is a population-based cohort of a predominantly white population in west of Sydney, Australia. At baseline (1992-94), 3,654 permanent residents aged 49 years or older participated (participation rate of 82.4%[11]. During 1997-99 (BMES II A), 2,335 participants (75.1% of survivors) returned for examinations after 5 years. During 1999-2000, 1,174 (85.2%) new participants took part in an Extension Study of the BMES (BMES IIB). BMES cross-section II thus includes BMES IIA (66.5%) and BMES IIB (33.5%) participants (n=3,509)[12]. From the BMES cross section II who had blood samples collected, DNA was extracted for 3,189 (90.1 %) participants. Over 98% of BMES participants were European ancestry. All BMES examinations were approved by the Human Ethics Committees of the Western Sydney Area Health Service and University of Sydney. Signed informed consent was obtained from participants at each examination.

Acknowledgements BMES : **BMES** was supported by the Australian National Health & Medical Research Council (NH&MRC), Canberra Australia (974159, 211069, 457349, 512423, 475604, 529912); the Centre for Clinical Research Excellence in Translational Clinical Research in Eye Diseases; NH&MRC research fellowships (358702, 632909 to J.J.W, 1028444, 1138585 to P.N.B.); and the Wellcome Trust, UK as part of Wellcome Trust Case Control Consortium 2 (A. Viswanathan, P. McGuffin, P. Mitchell, F. Topouzis, P. Foster) for genotyping costs of the entire BMES population (085475B08Z, 08547508Z, 076113). The Centre for Eye Research Australia receives Operational Infrastructure Support from the Victorian government. BMES acknowledges Jie Jin Wang and Elena Rochtchina from the Centre for Vision Research, Department of Ophthalmology and Westmead Millennium Institute University of Sydney (NSW Australia); John Attia, Rodney Scott, Elizabeth G. Holliday from the University of Newcastle (Newcastle, NSW Australia); Jing Xie from the Centre for Eye Research Australia, University of Melbourne; Michael T. Inouye, Medical Systems Biology, Department of Pathology & Department of Microbiology & Immunology, University of Melbourne (Victoria, Australia); Ananth Viswanathan, Moorfields Eye Hospital (London, UK); Paul J. Foster, NIHR Biomedical Research Centre for Ophthalmology, UCL Institute of Ophthalmology & Moorfields Eye Hospital (London); Peter McGuffin, MRC Social Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, King's College (London, United Kingdom); Fotis Topouzis, Department of Ophthalmology, School of Medicine, Aristotle University of Thessaloniki, AHEPA Hospital (Thessaloniki, Greece); Xueling Sim, National University of Singapore; members of the Wellcome Trust Case Control Consortium 2.

CROATIA-Korčula Study

The CROATIA-Korčula study, Croatia, is a population-based, cross-sectional study in the island of Korčula that includes a total of 969 adult examinees, aged 18-98 (mean=56.3), and most (N=930) underwent a complete eye examination[13]. The study received approval from relevant ethics committees in Scotland and Croatia and followed the tenets of the Declaration of Helsinki.

CROATIA-Split Study

The CROATIA-Split study, Croatia, is a population-based, cross-sectional study in the Dalmatian City of Split that includes 1000 examinees aged 18-95. The study received approval from relevant ethics committees in Scotland and Croatia and followed the tenets of the Declaration of Helsinki.

CROATIA-Vis Study

The CROATIA-Vis study, Croatia, is a population-based, cross-sectional study in the island of Vis including adult participants, aged 18–93 years (mean = 56), a subset of which (N=640) underwent a complete eye examination in summer 2007 and provided their ophthalmologic history[13]. The study received approval from relevant ethics committees in Scotland and Croatia and followed the tenets of the Declaration of Helsinki.

Acknowledgements CROATIA-Korčula, CROATIA-Split and CROATIA-Vis studies: The **CROATIA** studies were funded by grants from the Medical Research Council (UK), from the Republic of Croatia Ministry of Science, Education and Sports (108-1080315-0302; 216-1080315-0302) and the Croatian Science Foundation (8875); and the CROATIA-Korčula genotyping was funded by the European Union framework program 6 project EUROSPAN (LSHGCT2006018947). The CROATIA studies acknowledges Dr. Biljana Andrijević Derk, Valentina Lacmanović Lončar, Krešimir Mandić, Antonija Mandić, Ivan Škegro, Jasna Pavičić Astaloš, Ivana Merc, Miljenka Martinović, Petra Kralj, Tamara Knežević and Katja Barać-Juretić as well as the recruitment team from the Croatian Centre for Global Health, University of Split and the Institute of Anthropological Research in Zagreb for the ophthalmological data collection; Peter Lichner and the Helmholtz Zentrum Munchen (Munich, Germany), AROS Applied Biotechnology, Aarhus, Denmark and the Wellcome Trust Clinical facility (Edinburgh, United Kingdom) for SNP array genotyping; genetic analyses were supported by the MRC HGU "QTL in Health and Disease" core programme.

Diabetes Control and Complications Trial (DCCT)

DCCT (1982-1993) was a multi-center randomized clinical trial to compare the effectiveness of intensive (≥3 daily insulin injections or insulin pump) and conventional (<3 daily insulin injections) diabetic treatments at the time in preventing development and progression of microvascular complications of type 1 diabetes. Subjective refraction was measured following the standard protocols using a letter chart at 10 to 20 feet, at baseline visit and annually thereafter during DCCT. Refraction measurement was attempted at 1 meter for the subjects with poor visual acuity. In these cases the 4 meter refraction was estimated by subtracting +0.75 sphere from the 1 m measurement[14]. In the current study measurements at baseline were analyzed.

Acknowledgements DCCT: A complete list of researchers in the DCCT/EDIC Research Group is presented in the Supplementary Material published online[15]. Industry contributors have had no role in the DCCT/EDIC study but have provided free or discounted supplies or equipment to support participants' adherence to the study: Abbott Diabetes Care (Alameda, CA), Animas (Westchester, PA), Bayer Diabetes

Care (North America Headquarters, Tarrytown, NY), Becton Dickinson (Franklin Lakes, NJ), Eli Lilly (Indianapolis, IN), Extend Nutrition (St. Louis, MO), Insulet Corporation (Bedford, MA) , LifeScan (Milpitas, CA), Medtronic Diabetes (Minneapolis, MN), Nipro Home Diagnostics (Ft. Lauderdale, FL), Nova Diabetes Care (Billerica, MA), Omron (Shelton, CT), Perrigo Diabetes Care (Allegan, MI), Roche Diabetes Care (Indianapolis, IN) , and Sanofi-Aventis (Bridgewater, NJ). GWAS results from DCCT/EDIC will be made available through dbGaP. The DCCT/EDIC has been supported by cooperative agreement grants (1982-1993, 2012-2017), and contracts (1982-2012) with the Division of Diabetes Endocrinology and Metabolic Diseases of the National Institute of Diabetes and Digestive and Kidney Disease (current grant numbers U01 DK094176 and U01 DK094157), and through support by the National Eye Institute, the National Institute of Neurologic Disorders and Stroke, the General Clinical Research Centers Program (1993-2007), and Clinical Translational Science Center Program (2006-present), Bethesda, Maryland, USA. Trial Registration: clinicaltrials.gov NCT00360815 and NCT00360893. Additional support for this DCCT/EDIC collaborative study was provided by JDRF grant # 17-2013-9.

Estonian Genome Center, University of Tartu (EGCUT)

The Estonian cohort is from the population-based biobank of the Estonian Genome Center of the University of Tartu (EGCUT). The whole project is conducted according to the Estonian Gene Research Act and all participants have signed the broad informed consent (http://www.biobank.ee[16]). The current cohort size is over 51,515, from 18 years of age and up, which reflects closely the age distribution in the adult Estonian population. Subjects are recruited by the general practitioners (GP), physicians in the hospitals,and special recruitment offices of the EGCUT. They were randomly selected from the individuals visiting GP offices or hospitals. Computer Assisted Personal interviews were conducted by primary care providers and nurses during 1-2 hours at a doctor's office to collect information that includes personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women's health, quality of life) etc. Anthropometric and physiological measurements were also taken. All diseases are defined according to the ICD10 coding[17].

Acknowledgements EGCUT: **EGCUT** was supported by European Union H2020 grant 692145, Est.RC grant IUT20-60 and EU Project No. 2014-2020.4.01.15-0012. They received financing by FP7 grants (201413, 245536); Estonian Government (SF0180142s08); and the European Union through the European Regional Development Fund, in the frame of Centre of Excellence in Genomics and Estonian Research Infrastructure's Roadmap; EFSD grant; and the University of Tartu (SP1GVARENG). EGCUT acknowledges Mr. T. Esko, Mr. V. Soo, Ms. M-L. Tammesoo.

European Prospective Investigation into Cancer (EPIC-Norfolk)

The European Prospective Investigation into Cancer (**EPIC**) study is a pan-European prospective cohort study designed to investigate the aetiology of major chronic diseases[18]. EPIC-Norfolk, one of the UK arms of EPIC, recruited and examined 25,639 participants aged 40-79 years between 1993 and 1997 for the baseline examination[19]. Recruitment was via general practices in the city of Norwich and the surrounding small towns and rural areas, and methods have been described in detail previously[20]. Since virtually all residents in the UK are registered with a general practitioner through the National Health Service, general practice lists serve as population registers. Ophthalmic assessment formed part of the third health examination and this has been termed the EPIC-Norfolk Eye Study[21]. In total, 8,623 participants were seen for the ophthalmic examination, between 2004 and 2011. Refractive error was measured using a Humphrey Auto-Refractor 500 (Humphrey Instruments, San Leandro, California, USA). Genotyping was undertaken using the Affymetrix GeneChip Human Mapping 500K Array Set. Data were pre-phased with SHAPEIT version 2 and imputed to the March 2012 build of the 1000 Genomes project using IMPUTE version 2.2.2. The EPIC-Norfolk Eye Study was carried out following the principles of the Declaration of Helsinki and the Research Governance Framework for Health and Social Care. The study was approved by the Norfolk Local Research Ethics Committee (05/Q0101/191) and East Norfolk & Waveney NHS Research Governance Committee (2005EC07L). All participants gave written, informed consent.

Acknowledgements EPIC-Norfolk: EPIC-Norfolk infrastructure and core functions are supported by grants from the Medical Research Council (G1000143) and Cancer Research UK (C864/A14136). The clinic for the third health examination was funded by Research into Ageing (262). Mr Khawaja is a Wellcome Trust Clinical Research Fellow. Miss Chan is a joint Medical Research Council/Royal College of Ophthalmologists Research Fellow, and received additional support from the International Glaucoma Association. Professor Foster has received additional support from the Richard Desmond Charitable Trust (via Fight for Sight) and the Department for Health through the award made by the National Institute for Health Research to Moorfields Eye Hospital and the UCL Institute of Ophthalmology for a specialist Biomedical Research Centre for Ophthalmology.

Erasmus Rucphen Family Study (ERF)

The Erasmus Rucphen Family (ERF) Study is a family-based cohort in a genetically isolated population in the southwest of the Netherlands with over 3,000 participants aged between 18 and 86 years. Cross-sectional examination took place between 2002 and 2005. The rationale and study design of this study have been described elsewhere[22,23]. Cross-sectional examination took place between 2002 and 2005, including a non-dilated automated measurement of refractive error using a Topcon RM-A2000 autorefractor. All measurements in these studies were conducted after the Medical Ethics Committee of the Erasmus University had approved the study protocols and all participants had given a written informed consent in accordance with the Declaration of Helsinki.

Acknowledgements ERF: see Rotterdam Study.

Fuchs' Endothelial Corneal Dystrophy Controls (FECD)

We utilized control subjects who were part of a larger study on the genetics of Fuchs' Endothelial Corneal Dystrophy (FECD)[10]. All control subjects were of European descent and were at least 60 years of age and matched according to age, gender, and ancestry to the enrolled index cases. To qualify, each control subject was required to be grade 0 on the FECD grading scale, have no family history of a possibly inherited corneal disorder (eg, FECD, keratoconus, stromal dystrophy), and have normal corneas with no abnormalities on slit-lamp examination apart from certain conditions judged not to affect FECD[10]. Subjects were excluded from participation as controls if they displayed any signs of corneal dystrophy or degeneration or had previous/active interstitial keratitis or anterior uveitis, or active/previous infectious keratitis, or vascularization of the epithelium and/or stroma. Subjects were also excluded if they had previously undergone bilateral corneal surgery or had experienced perforating corneal trauma resulting in scarring. Measurements of refractive error, central corneal thickness and absence of FECD were obtained at baseline, along with recorded age, and gender. This work was performed in accordance with the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants. Data were collected under multi-center Institutional Review Board (IRB) approval.

Acknowledgements FECD: see BMES

Finnish Twin Study on Aging (FITSA)

Finnish Twin Study on Aging (FITSA) is a study of genetic and environmental effects on the disablement process in older female twins[24]. The study cohort of 13 888 adult twin pairs started in 1975. Altogether 103 MZ and 114 DZ twin pairs (424 individuals, all women of European ancestry) aged 63-76 years living in Finland took part in multiple laboratory examinations in 2000 and 2003, and responded in

questionnaires in 2011. Before the examinations, the subjects provided a written informed consent according to the Declaration of Helsinki. The study protocol was approved by the ethics committee of the Central Hospital District of Central Finland.

Acknowledgements FITSA: FITSA was supported by ENGAGE (FP7-HEALTH-F4-2007, 201413); European Union through the GENOMEUTWIN project (QLG2-CT-2002-01254); the Academy of Finland Center of Excellence in Complex Disease Genetics (213506, 129680); the Academy of Finland Ageing Programme; and the Finnish Ministry of Culture and Education and University of Jyväskylä. For FITSA the contributions of Emmi Tikkanen, Samuli Ripatti, Markku Kauppinen, Taina Rantanen and Jaakko Kaprio are acknowledged.

Framingham Eye Study

The Framingham Eye Study[25] (FES) was nested within the Framingham Heart Study (FHS, http://www.framinghamheartstudy.org), which began its first round of extensive physical examinations in 1948 by recruiting 5,209 men and women from the town of Framingham, MA, USA. Surviving participants from the original cohort returned for biennial exams, which continue to the present. A total of 2675 FHS participants were also examined as part of the FES between 1973 and 1975. The FES was designed to evaluate ocular characteristics of examinees such as: senile cataract; age-related macular disease; glaucoma; and retinopathy. Between 1989 and 1991, 1603 offspring of original cohort participants also received ocular examinations[26]. The analyses in the current study are limited 1497 (42.5% men) participants from both the original and the offspring cohorts for whom genotype data were available. Most individuals in this analysis set are unrelated but a small number of related pairs remain. All data--including refractive error, demographics and genotypes--were retrieved from the database of Genotypes and Phenotypes (dbGaP, http://www.ncbi.nlm.nih.gov/gap) after approval for controlled access to individual-level data. All study protocols are in compliance with the World Medical Association Declaration of Helsinki. Since 1971, written consent has been obtained from participants before each examination. The research protocols of the Framingham Heart Study are reviewed annually by the Institutional Review Board of the Boston University Medical Center and by the Observational Studies Monitoring Board of the National Heart, Lung and Blood Institute.

Acknowledgements Framingham Eye Study: **Framingham Eye Study** was supported by NEI (N01EY22112, N01EY92109); the National Heart, Lung, and Blood Institute (N02HL64278) for SHARe genotyping; Boston University (N01HC25195); and by intramural funds of the National Human Genome Research Institute, NIH, USA (to R.W. and J.E.B.W.).

Gutenberg Health Study (GHS1, GHS2)

The **Gutenberg Health Study** is a population-based, prospective, observational cohort study in mid-western Germany that includes consecutive follow-ups every five years. The primary study aim is to evaluate and improve cardiovascular risk stratification and the general health status of the population. The baseline examination included a total of 15,010 participants aged 35 to 74 years and took place from 2007 to 2012. The participants were randomly drawn and equally stratified for sex, residence (urban or rural) and for each decade of age. Exclusion criteria were the following: insufficient knowledge of German and physical or mental inability to participate in the examinations in the study center. The ophthalmic examination was based on standard operating procedures and included without limitation autorefraction and visual acuity testing (Humphrey® Automated Refractor/Keratometer (HARK) 599™, Carl Zeiss Meditec AG, Jena, Germany). The study protocol and study documents were approved by the local ethics committee of the Medical Chamber of Rhineland-Palatinate, Germany (reference no. 837.020.07; original vote: 22.3.2007, latest update: 20.10.2015). According to the tenets of the Declaration of Helsinki, written informed consent was obtained from all participants prior to their entry into the study.

Acknowledgements Gutenberg Health Study (GHS1, GHS2): The Gutenberg Health Study is funded through the government of Rhineland-Palatinate („Stiftung Rheinland-Pfalz für Innovation“, contract AZ 961-386261/733), the research programs “Wissen schafft Zukunft” and “Center for Translational Vascular Biology (CTVB)” of the Johannes Gutenberg-University of Mainz, the National Genome Network "NGFNplus" by the Federal Ministry of Education and Research, Germany (A301GS0833) and its contract with Boehringer Ingelheim, PHILIPS Medical Systems and Novartis Pharma. We thank all study participants for their willingness to provide data for this research project and we are indebted to all coworkers for their enthusiastic commitment.

KORA

KORA ("Kooperative Gesundheitsforschung in der Region Augsburg" which translates as “Cooperative Health Research in the Region of Augsburg”) is a population based study of adults randomly selected from 430,000 inhabitants living in Augsburg and 16 surrounding counties in Germany[27-29]. The collection was done in 4 separate groups from 1984-2001 (S1-S4). All survey participants are residents of German nationality identified through the registration office. In the KORA S3 and S4 studies 4,856 and 4,261 subjects have been examined implying response rates of 75% and 67%, respectively. 3,006 subjects participated in a 10-year follow-up examination of S3 in 2004/05 (KORA F3), and 3080 of S4 in 2006/2008 (KORA F4). The age range of the participants was 25 to 74 years at recruitment. The study

was approved by the local ethics committee. Written informed consent was obtained from all participants before enrollment in accordance with the Declaration of Helsinki.

Acknowledgements KORA: **KORA** was financed by the Helmholtz Center Munich, German Research Center for Environmental Health; the German Federal Ministry of Education and Research; the State of Bavaria; the German National Genome Research Network (NGFN-2 and NGFNPlus) (01GS0823); Munich Center of Health Sciences as part of LMUinnovativ; the genotyping was carried out by the Center for Inherited Disease Research, Baltimore, MD, and was supported by the National Eye Institute (R01 EY020483 to D.S.). Some of the analyses were supported by the Intramural Research Program of the National Human Genome Research Insitute, NIH, USA.

OGP Ogliastra Genetic Park, Talana study (OGP Talana)

A cross-sectional ophthalmic study was performed in Talana, Perdasdefogu and Urzulei within the Ogliastra Project, a large epidemiological survey conducted in a geographically, culturally and genetically isolated population living in an eastern-central region of Sardinia[30]. In Talana the study was carried out between October 2001 and October 2002 and adhered to the tenets of the declaration of Helsinki. Talana is a village situated at an altitude of 700 m above sea level in one of the most secluded areas of Sardinia, Ogliastra; it has about 1200 inhabitants and, importantly, archival records are available from 1589 and genealogical trees have been reconstructed from 1640. 789 volunteers gave their written informed consent and were invited to the local medical centre, which was equipped with a complete set of ophthalmic instruments for this survey. Participants underwent a complete eye examination including visual acuity (Snellen charts, 5 m) and refraction status assessment (autorefractor RK-8100 Topcon, Tokyo, Japan)

Acknowledgements OGP Ogliastra Genetic Park, Talana study (OGP Talana): **OGP Talana** was supported by grants from the Italian Ministry of Education, University and Research (5571DSPAR2002, 718Ric2005). OGP Talana thanks the Ogliastra population and the municipal administrators for their collaboration to the project and for economic and logistic support.

Orkney Complex Disease Study (ORCADES)

The Orkney Complex Disease Study (ORCADES) is a population-based, cross-sectional study in the Scottish archipelago of Orkney, including 1,285 individuals with eye measurements. The study received approval from relevant ethics committees in Scotland and followed the tenets of the Declaration of Helsinki.

Acknowledgements ORCADES: **ORCADES** recruitment and genotyping were supported by the Chief Scientist Office of the Scottish Government, the Royal Society, the UK Medical Research Council Human Genetics Unit and the European Union framework program 6 EUROSPAN project (LSHGCT2006018947). ORCADES acknowledges the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, in particular Margaret Pratt who performed the eye measurements, as well as the administrative team in Edinburgh University, the Wellcome Trust Clinical facility (Edinburgh, United Kingdom) for DNA extraction, Peter Lichner and the Helmholtz Zentrum Munchen (Munich, Germany) for genotyping, and Mirna Kirin, Pau Navarro and Peter Joshi for the genetic data imputation. Genetic analyses were supported by the MRC HGU "QTL in Health and Disease" core programme.

Rotterdam Study (RS1, RS2, RS3)

The Rotterdam Study is a prospective population-based cohort study in the elderly living in Ommoord, a suburb of Rotterdam, the Netherlands. Details of the study are described elsewhere[31]. In brief, the Rotterdam Study consists of 3 independent cohorts: RS1, RS2, and RS3. For the current analysis, 5,328 residents aged 55 years and older were included from RS1, 2,009 participants aged 55 and older from RS2, and 1,970 aged 45 and older from RS 3. 99% of subjects were of European ancestry. Participants underwent multiple physical examinations with regular intervals from 1991 to present, including a non-dilated automated measurement of refractive error using a Topcon RM-A2000 autorefractor. All measurements in RS-1–3 were conducted after the Medical Ethics Committee of the Erasmus University had approved the study protocols and all participants had given a written informed consent in accordance with the Declaration of Helsinki.

Acknowledgements Rotterdam Study and ERF: **The Rotterdam Study** and **ERF** were supported by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant 648268), Netherlands Organisation for Scientific Research (NWO, grant 91815655), Erasmus Medical Center and Erasmus University, Rotterdam, The Netherlands; Netherlands Organization for Health Research and Development (ZonMw); UitZicht; Netherlands Organisation for Scientific Research (NWO Veni 91617076 to V.J.M.V.); the Research Institute for Diseases in the Elderly; the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission (DG XII); the Municipality of Rotterdam; the Netherlands Genomics Initiative/NWO; Center for Medical Systems Biology of NGI; Lijf en Leven; M.D. Fonds; Henkes Stichting; Stichting Nederlands Oogheelkundig Onderzoek; Swart van Essen; Bevordering van Volkskracht; Blindenhulp; Landelijke Stichting voor Blinden en Slechtzienden; Rotterdamse Vereniging voor Blindenbelangen; Oogfonds; Algemene Nederlandse Vereniging ter Voorkoming van Blindheid; Stichting MaculaFonds;

Combined Ophthalmic Research Rotterdam; Rotterdamse Oogheelkundig Onderzoek Stichting; Erasmus MC Vriendenfonds, Topcon Europe; Novartis; Ada Hooghart, Corina Brussee, Riet Bernaerts-Biskop, Patricia van Hilten, Pascal Arp, Jeanette Vergeer, Marijn Verkerk; Sander Bervoets.

TEST/BATS

The Australian Twin Eye Study comprises participants examined as part of the Twins Eye Study in Tasmania or the Brisbane Adolescent Twins Study. Details of the study are described elsewhere[32]. Ethical approval was obtained from the Royal Victorian Eye and Ear Hospital, the University of Tasmania, the Australian Twin Registry and the Queensland Institute of Medical Research.

Acknowledgements TEST/BATS: **TEST** and **BATS** (Australian Twins) were supported by an Australian National Health and Medical Research Council (NHMRC) Enabling Grant (2004-2009, 350415, 2005-2007); Clifford Craig Medical Research Trust; Ophthalmic Research Institute of Australia; American Health Assistance Foundation; Peggy and Leslie Cranbourne Foundation; Foundation for Children; Jack Brockhoff Foundation; National Institutes of Health/National Eye Institute (RO1EY01824601 (2007-2010)); Pfizer Australia Senior Research Fellowship (to D.A.M.); and Australian NHMRC Career Development Award (to S.M.). Genotyping was funded by an NHMRC Medical Genomics Grant; US National Institutes of Health/National Eye Institute (1RO1EY018246), Australian sample imputation analyses were carried out on the Genetic Cluster Computer which is financially supported by the Netherlands Scientific Organization (NWO48005003). Australian Twins thanks Grant Montgomery, Scott Gordon, Dale Nyholt, Sarah Medland, Brian McEvoy, Margaret Wright, Anjali Henders, Megan Campbell for ascertaining and processing genotyping data; Jane MacKinnon, Shayne Brown, Lisa Kearns, Jonathan Ruddle, Paul Sanfilippo, Sandra Staffieri, Olivia Bigault, Colleen Wilkinson, Yaling Ma, Julie Barbour for assisting with clinical examinations; and Dr Camilla Day and staff at the Center for Inherited Disease Research.

TwinsUK

The TwinsUK adult twin registry based at St. Thomas' Hospital in London is a volunteer cohort of over 10,000 twins from the general population[33]. Twins largely volunteered unaware of the eye studies, gave fully informed consent under a protocol reviewed by the St. Thomas' Hospital Local Research Ethics Committee.

Acknowledgements TwinsUK: **TwinsUK** received funding from the Wellcome Trust; the European Union MyEuropia Marie Curie Research Training Network; Guide Dogs for the Blind Association; the European

Community's FP7 (HEALTHF22008201865GEFOS); ENGAGE (HEALTHF42007201413); the FP-5 GenomEUtwin Project (QLG2CT200201254); US National Institutes of Health/National Eye Institute (1RO1EY018246); NIH Center for Inherited Disease Research; the National Institute for Health Research comprehensive Biomedical Research Centre award to Guy's and St. Thomas' National Health Service Foundation Trust partnering with King's College London. P.G.H. is the recipient of a Fight for Sight ECI award. We acknowledge the contribution of Drs Toby Andrew, Margarida Lopes, Samantha Fahy and Diana Kozareva.

Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR)

WESDR is an observational cohort study of diabetes complications (1979-2007)[34]. Subjective refraction, measured following standard protocols at first visit, was analyzed in the current study (n=589).

Acknowledgements WESDR: **WESDR** was supported by NEI (grants R01EY03083 and EY016379) and a Research to Prevent Blindness Senior Scientific Investigator Award.

Young Finns Study (YFS)

The YFS cohort is a Finnish longitudinal population study sample on the evolution of cardiovascular risk factors from childhood to adulthood[35]. The first cross-sectional study was conducted in the year 1980 in five different centers. It included 3,596 participants in the age groups of 3, 6, 9, 12, 15, and 18, who were randomly chosen from the national population register. After the baseline in 1980 these subjects have been re-examined in 1983 and 1986 as young individuals, and in 2001, 2007 and 2011 as older individuals. For the current analysis a subsample from the newest (2011) follow-up was used from four centers (N=1479) where the refractive error measurements data from both eyes were available.

This study was carried out in accordance with the recommendations of the Declaration of Helsinki. All participants provided written informed consent and the study protocol was approved by the Ethics Committee.

Acknowledgements Young Finns Study (YFS): The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Competitive State Research Financing of the Expert Responsibility area of Kuopio, Tampere and Turku University Hospitals (grant X51001); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation for Cardiovascular Research ; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation;

Yrjö Jahnsson Foundation; Signe and Ane Gyllenberg Foundation; and Diabetes Research Foundation of Finnish Diabetes Association.

Beijing Eye Study (BES)

The BES is a population-based cohort of Han Chinese in the rural region and in the urban region of Beijing in North China. The Medical Ethics Committee of the Beijing Tongren Hospital approved the study protocol and all participants gave informed consent, according to the Declaration of Helsinki. At baseline (2001), 4439 individuals out of 5324 eligible individuals aged 40 years or older participated (response rate: 83.4%). In the years 2006 and 2011, the study was repeated by re-inviting all participants from the survey from 2001 to be re-examined. Out of the 4439 subjects examined in 2001, 3251 (73.2%) subjects returned for the follow-up examination in 2006, and 2695 (60.7%) subjects returned for the follow-up examination in 2011.

Acknowledgements Beijing Eye Study: **Beijing Eye Study** was supported by National Natural Science Foundation of China (grant 81170890).

Nagahama

Nagahama Prospective Cohort for Comprehensive Human Bioscience (the Nagahama Study) is a community-based cohort consisted of 9,804 healthy Japanese volunteers recruited between 2008 and 2010 from the general population of Nagahama City in Japan. Community residents from 30–74 years of age, living independently and without physical impairment or dysfunction were eligible. The Kyoto University Graduate School and Faculty of Medicine Ethics Committee and the Nagahama Municipal Review Board of Personal Information Protection approved the study protocol and procedures used to obtain informed consent. All the study procedures adhered to the tenets of the Declaration of Helsinki. All participants were fully informed about the purpose and procedures of the study, and written consent was obtained from each subject.

Acknowledgements Nagahama: Nagaham Study was financially supported by Comprehensive Research on Aging and Health Science Research Grants for Dementia R&D from Japan Agency for Medical Research and Development (AMED) and the Centre of Innovation Program, the Global University Project from Japan Science and Technology Agency.

Singapore Studies

All Singapore studies adhere to the Declaration of Helsinki. Ethics approvals have been obtained from the Institutional Review Boards of the Singapore Eye Research Institute, Singapore General hospital, National University of Singapore and National Healthcare Group, Singapore. In all cohorts, participants provided written, informed consent at the recruitment into the studies.

Singapore Prospective Study Program (SP2)

Samples of SP2 were from a revisit of two previously conducted population-based surveys carried out in Singapore between 1992 and 1998, including the National Health Survey 1992 and the National Health Survey 1998[36]. These studies comprise random samplings of individuals stratified by ancestry from the entire Singapore population. A total of 8266 subjects were invited in this follow-up survey and 6301 (76.1% response rate) subjects completed the questionnaire, of which 4056 (64.4% of those who completed the questionnaire) also attended the health examination and donated blood specimens.

Acknowledgements Singapore Prospective Study Program (SP2): See "Acknowledgements Singapore Studies"

Singapore Malay Eye Study (SiMES)

SiMES is a population-based prevalence survey of Malay adults aged 40 to 79 years living in Singapore that was conducted between August of 2004 and June of 2006[37]. From a Ministry of Home Affairs random sample of 16,069 Malay adults in the Southwestern area, an age-stratified random sampling strategy was used in selecting 1400 from each decade from age 40 years onward (40–49, 50–59, 60–69, and 70–79 years).The 4,168 eligible participants from the sampling frame, while 3280 (78.7%) participated.

Acknowledgements Singapore Malay Eye Study (SiMES)**:** *See "Acknowledgements Singapore Studies"*

Singapore Indian Eye Study (SINDI)

SINDI is a population-based survey of major eye diseases[38] in ethnic Indians aged 40 to 80 years living in the South-Western part of Singapore and was conducted from August 2007 to December 2009. In brief, 4,497 Indian adults were eligible and 3,400 participated.

Acknowledgements Singapore Indian Eye Study (SINDI): *See "Acknowledgements Singapore Studies"*

Singapore Chinese Eye Study (SCES)

Similar to SINDI, the Singapore Chinese Eye Study (SCES) is a population-based cross-sectional study of eye diseases in Chinese adults 40 years of age or older residing in the southwestern part of Singapore. The methodology of the SCES study has been described in details previously. Between 2009 and 2011, 3,353 (72.8%) of 4,605 eligible individuals underwent a comprehensive ophthalmologic examination, using the same protocol as SINDI[37].

Acknowledgements Singapore Chinese Eye Study (SCES): see "Acknowledgements Singapore Studies" STARS

Acknowledgements Singapore Studies: The Singapore studies (**SP2, SIMES, SINDI, SCES, STARS**) were supported by the National Medical Research Council, Singapore (NMRC 0796/2003, NMRC 1176/2008, STaR/0003/2008; CG/SERI/2010), Biomedical Research Council, Singapore (06/1/21/19/466, 09/1/35/19/616 and 08/1/35/19/550). The Singapore Tissue Network and the Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore provided services. National supercomputing centre (NSCC) provided high performance computing resources to support GWAS analysis.

Acknowledgements UK Biobank: This research was facilitated by data from the UK Biobank Resource. UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council (UK), Department of Health (UK), Scottish Government, and Northwest Regional Development Agency. It also had funding from the Welsh Assembly Government, British Heart Foundation, and Diabetes UK. The eye and vision dataset has been developed with additional funding from The NIHR Biomedical Research Centre at Moorfields Eye Hospital and the UCL Institute of Ophthalmology, Fight for Sight charity (UK), Moorfields Eye Charity (UK), The Macula Society (UK), The International Glaucoma Association (UK) and Alcon Research Institute (USA). This research has been conducted using the UK Biobank Resource (applications #17351 and #17615).

Acknowledgements expression study Young TL et al: This study was funded by the National Institutes of Health R01 EY014685, the Research to Prevent Blindness Inc. Lew R. Wasserman Award, and the University of Wisconsin Centennial Scholars Award (T.L. Young).

2: methods per subfields of prioritization

Internal replication of index genetic variants in the individual cohort GWAS'es (CREAM-ASN, CREAM-EUR and 23andMe)

Internal replication of the index genetic variants from Stage 3 were checked in Stage 1 and 2 using the Bonferroni corrected p-value of $0.05/(140*3) = 1.19x10^{-4}$.

Evidence for eQTL

We used FUMA[39] for eQTL look-ups which was based on eQTLs and gene expression used in the pipeline were obtained from GTEx v6[40], eQTLs of blood cells from the Blood eQTL Browser[41] (cis-eQTLs with FDR ≤ 0.05), eQTLs of blood cells in Dutch population from the BIOS QTL Browser (cis-eQTLs on gene-level with FDR ≤ 0.05), eQTLs of 10 brain regions from BRAINEAC ((http://www.braineac.org/,; sis-eQTLs with nominal *P* value < 0.05) and additional extensive look-ups in GtEx.

Evidence of expression in the eye in developmental and adult ocular tissues

Bergen AA et al:

Gene expression of myopia candidate genes in laser dissected, freshy frozen retinal tissues: human RPE, photoreceptors and choroid. Cellular expression was measured using validated 44K microarray data on an Agilent platform in multiple samples, and RNA levels were ranked in percentiles, with 100,00 indicating the highest expression, and 0 the lowest, according to methodology described by Booij et al (2009).

Young TL et al:

Expression of genes annotated to the index variants was studied in human ocular tissue using various methods: expression profiles in laser dissected freshly frozen RPE, photoreceptors and choroid of healthy human adult donor eyes[42]; whole-transcriptome expression analysis of macular and peripheral retina, choroid, and sclera from eight adult normal human ocular eyes[43]; and whole genome expression in retina, retinal pigment epithelium, choroid, sclera, optic nerve, and cornea from 15 normal fetal (24 weeks gestational age) and 6 adult donor eyes[44]. Human fetal and adult gene expression data for retina, RPE, choroid, sclera, optic nerve and cornea were obtained as described in Young TL, et al. (2013)[44]. In brief, 9 fetal eyes at 12-weeks' gestation and 6 fetal eyes at 24-weeks' gestation were obtained from Advanced Biosciences Resources (Alameda, CA, USA), while 6 adult eyes were obtained from the North Carolina

Eye Bank (Winston–Salem, North Carolina, USA). Whole globes with a 2mm equatorial incision were immersed in RNAlater (Qiagen, Hilden, Germany) shortly after collection to preserve RNA integrity, and shipped overnight on ice. The retina, RPE, choroid and scleral tissues were isolated at the posterior pole using round biopsy punches. Some fetal tissues, such as RPE and retina could not be separated, and were collected in combination. Central corneal samples were isolated using a biopsy punch, and optic nerve was collected using clean dissection scissors. Tissue samples were homogenized at 4°C in Ambion lysis buffer using a Bead Ruptor Tissue Homogenizer (Omni, Kennesaw, Georgia, USA) with 2.38 mm metal beads, and RNA was extracted using the mirVanaTM total RNA extraction kit (Ambion, Austin, Texas, USA) following the manufacturer's protocol. The RNA samples were labeled and amplified using the Illumina Total Prep kit (Ambion, Austin, Texas, USA), and hybridized to Illumina HumanHT-12 v4 Expression BeadChips (San Diego, California, USA). All protocols were performed following the manufacturer's recommendations. Twelve tissue samples were processed on each chip. Microarray data background noise was subtracted from the intensity values using Illumina's GenomeStudio software, exported and log2 transformed.

Presence of an eye phenotype in knock-out mice

The Mouse Genome Informatics database (MGI, www.informatics.jax.org) and the International Mouse Phenotyping Consortium (IMPC, http://www.mousephenotype.org) were checked for entries matching an eye-phenotype. Genes were listed and their human equivalents were looked up in NCBI (www.ncbi.nlm.nih.gov).

Presence of an eye phenotype in humans

All genes were checked in Online Mendelian Inheritance in Man (OMIM, http://omim.org) and DisGeNET (http://www.disgenet.org) for the involvement of a human ocular phenotype.

Presence of gene in a significant enriched functional pathway (DEPICT)

We first clumped the SNPs with *p*-value $<1x10^{-5}$ or lower from the meta-analysis of stage III using 500kb as physical distance threshold and an $R^2 > 0.1$ with PLINK, resulting in 534 clumps. Secondly, we performed gene set, cell type, and tissue enrichment analyses using DEPICT[45]. The Affinity Propagation tool[46-49] was used for clustering, and clusters were named by their 'representative' gene set, which was automatically chosen by the Affinity Propagation clustering method. The pairwise Pearson correlation between significant gene sets was calculated and then the AP algorithm was used to cluster similar

pathways into meta gene sets. Clusters were named by their representative gene set, which was automatically selected by the AP algorithm. In addition, correlation between the meta gene sets was calculated to create a network. The visualizations of the gene set enrichment analysis were created in Cytoscape[50].

Presence of gene in a significant canonical pathway (IPA)

IPA is a subscription based manually curated knowledge archive. Bioinformatic analysis according to Ingenuity protocol and data sets containing RefSeq identifiers were uploaded into the application IPA. Each RefSeq identifier was mapped to its corresponding human splicing variant in the Ingenuity® Knowledge Base. Canonical pathway analysis identified the 5 canonical pathways from the IPA library that were most significant to the data set. The significance of the association between the data set and the canonical pathway was measured in two ways: 1) A ratio of the number of molecules from the data set that map to the pathway divided by the total number of molecules that map to the canonical pathway. 2) Fisher's Exact test was used to calculate a P-value determining the probability that the association between the genes in the data set and the canonical pathway is explained by chance alone.

3: Dopamine pathway look-ups

To further investigate the association of genes playing a key role in the dopamine pathway, we we looked up the regions coding for dopamine receptors (*DRD1, DRD2, DRD3, DRD4, DRD5*), genes involved in synthesis & degradation (*COMT, DBH, DDC, MAOA, TH*) and transporters: (*SLC6A3/DAT, SLC6A4/SERT*) in the Stage 3 meta-analysis[51-53]. An overview of the results are provided in Supplementary Table 10 and Locus Zoom regional plots are provided in Supplementary Figure 12.

4: Phenotyping in CREAM and 23andMe

CREAM

Phenotyping of the CREAM cohorts has been described in detail previously[54]. Eligible participants underwent a complete ophthalmic examination, including a non-dilated measurement of refractive error for both eyes using a similar protocol (Supplementary Excel Table 1a). Inclusion criteria were individuals over the age of 25 years, of European or Asian descent, with available data on refractive error and with available genotype data. As previously described[54], exclusion criteria were all refraction altering conditions and ocular syndromes or systemic syndromes. Spherical equivalent (SE) was calculated according to the standard formula (SE = sphere + ½ cylinder); the mean SE of two eyes was used for analyses. When SE was only available for one eye, the SE of this eye was used.

23andMe

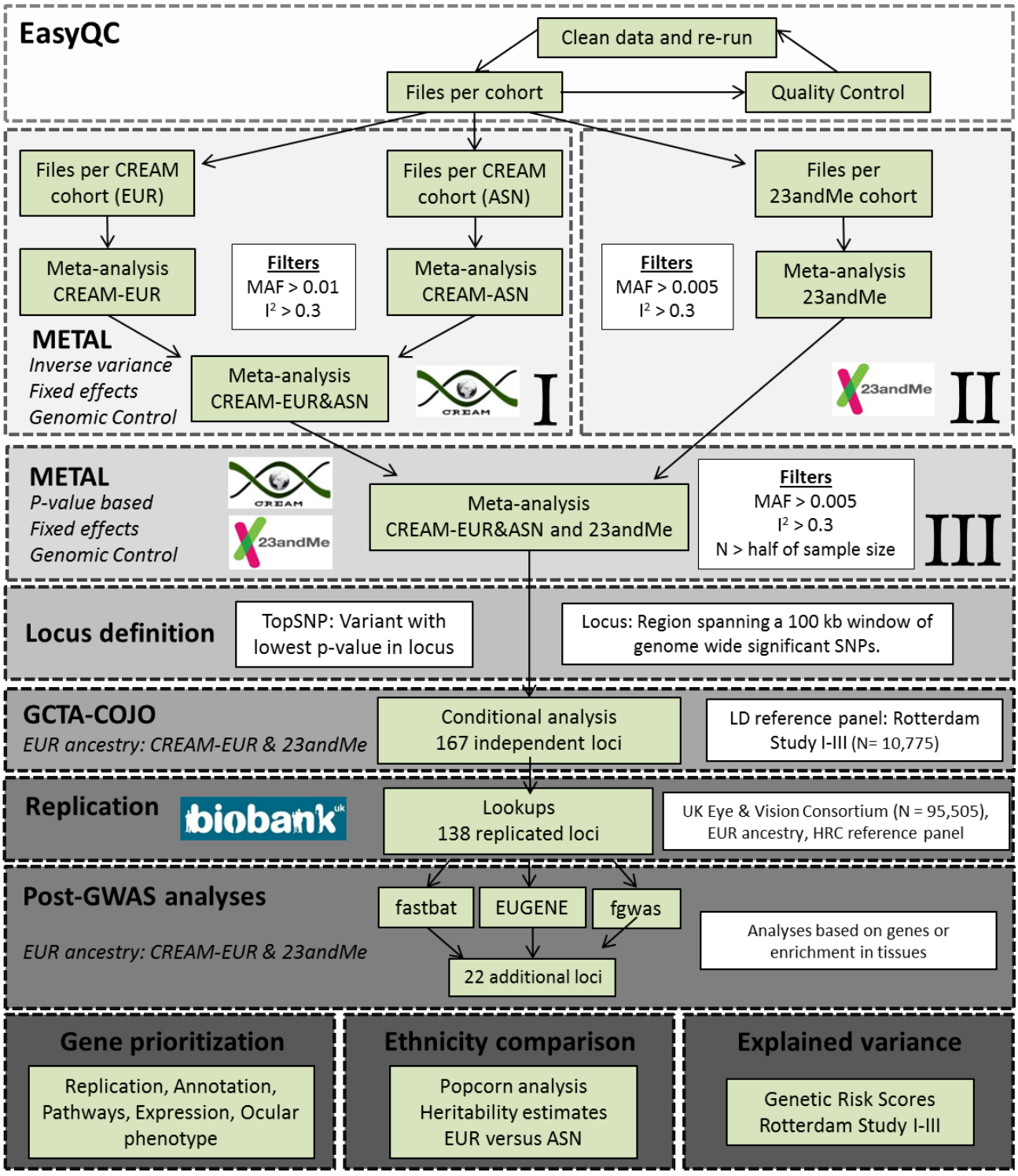
All participants were drawn from the customer database of 23andMe, Inc., a direct-to-consumer genetic testing company. Phenotyping of this cohort has been described in detail previously[55,56]. Participants were asked online whether they had myopia (yes or no), and at what age they were diagnosed with myopia. Inclusion criteria were individuals of European ancestry, age at diagnosis between five and 30 years of age, and available genotype data.

Participants provided informed consent and participated in the research online, under a protocol approved by the external AAHRPP-accredited IRB, Ethical & Independent Review Services (E&I Review).

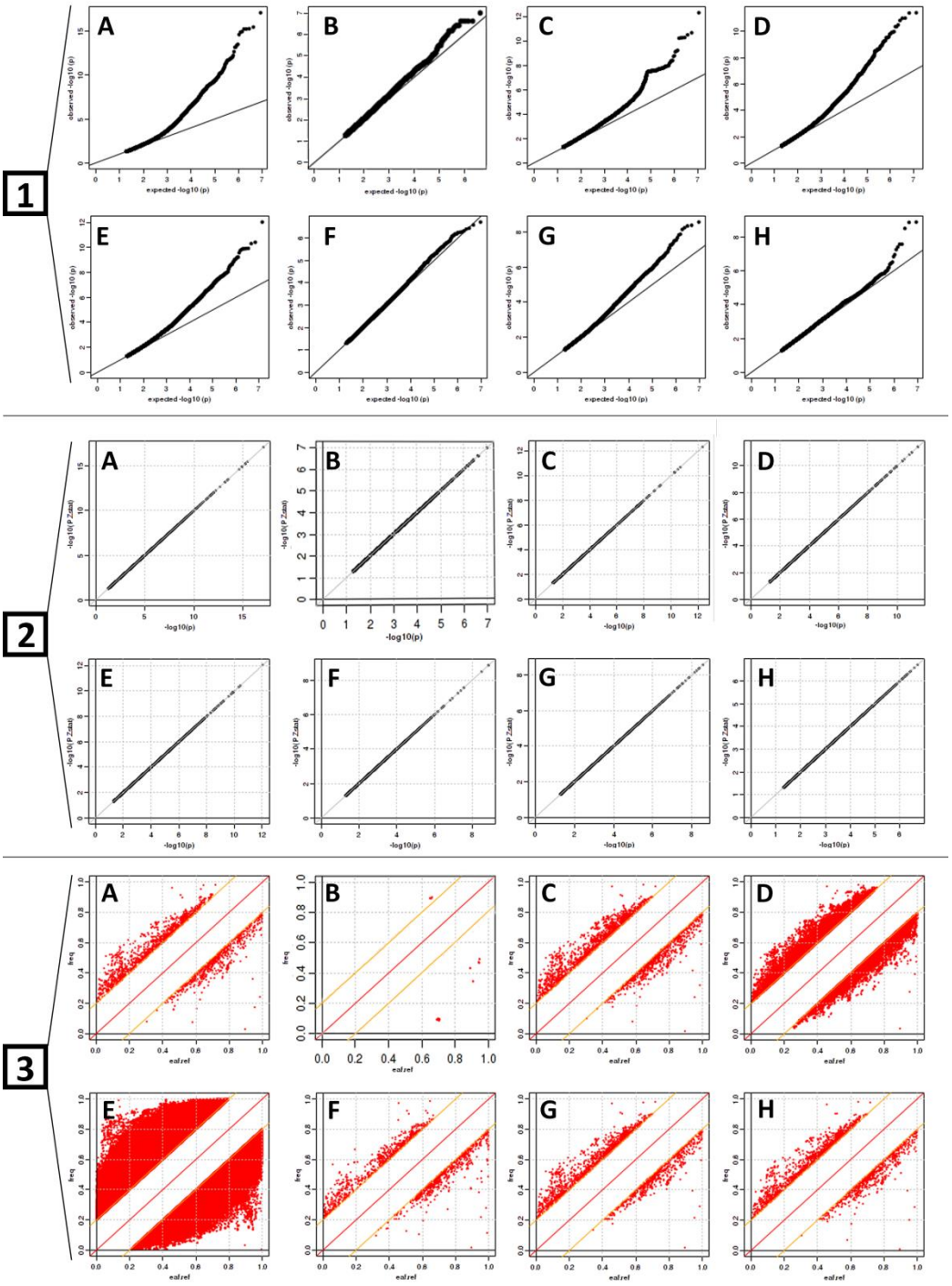
5: Gene-based test - fgwas

We looked at enrichment of 450 genomic annotations as implemented in fgwas with 5000 SNPs per window and . We derived the best annotations from genome-wide significant loci. We first considered annotations separately to see if they were individually significant. Some annotations were correlated and, hence, we built a model by adding terms sequentially in decreasing order of significance. We started with a model with two region based annotations (top one third of the distribution of gene density and bottom one third of the distribution of gene density). We then added two SNP based annotations (SNPs between 0 and 5kb from the transcription start site and SNPs between 5kb and 10kb from the transcription start site) and continued to add annotations until a maximum of 10 SNP based annotations were included in the model (adding further annotations is possible although we had problems with convergence within the fgwas software when a large number of annotations were added so we stopped adding further annotations at 10). We then applied the cross-validation approach implemented in fgwas to ensure no over-fitting in the final model. We used this final Bayesian model to derive a prior distribution for the remainder of the genome. We calculated the posterior probability of association based on the derived prior distribution. A posterior probability greater than 0.9 in this approach performed similarly to the traditional genome-wide significance threshold in genome-wide association studies (p-value$<5 \times 10^{-8}$) based on the analysis of previously published genome-wide association studies.

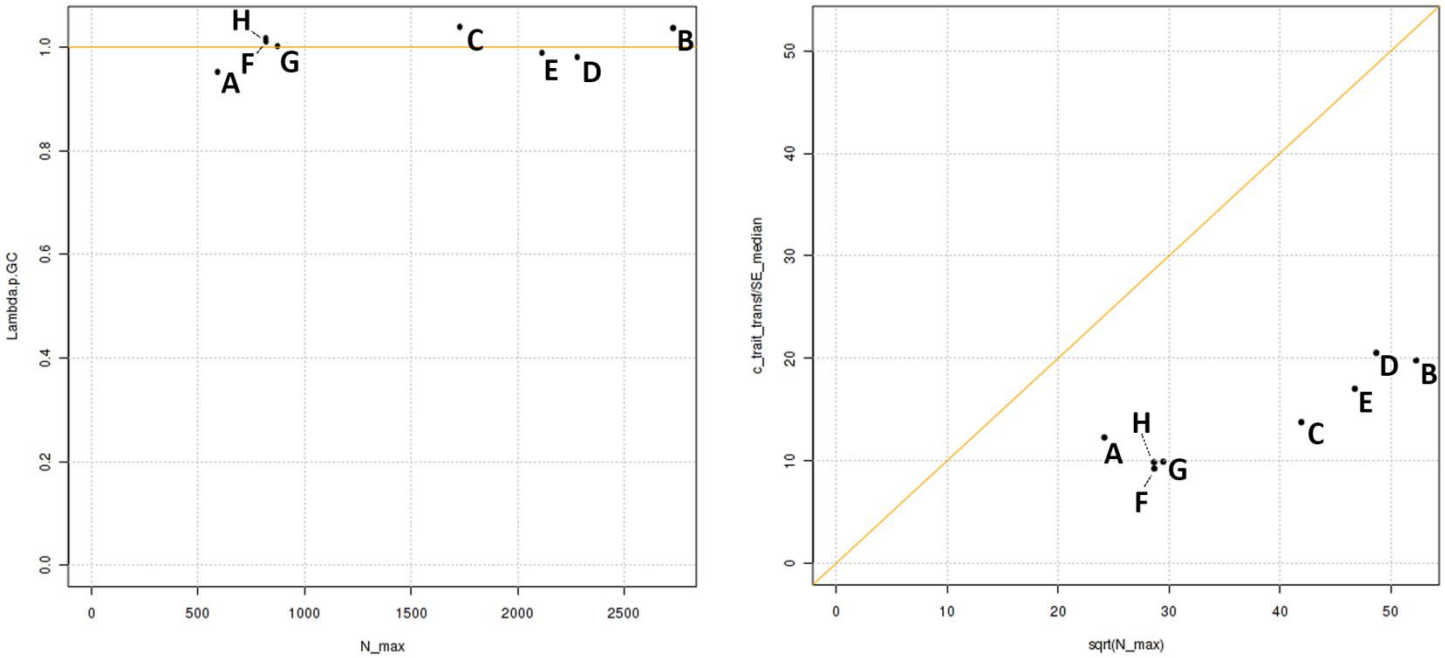
Supplementary Figure 1: Study design flow-chart



Supplementary Figure 2.1: EasyQC plots per cohort – CREAM-ASN



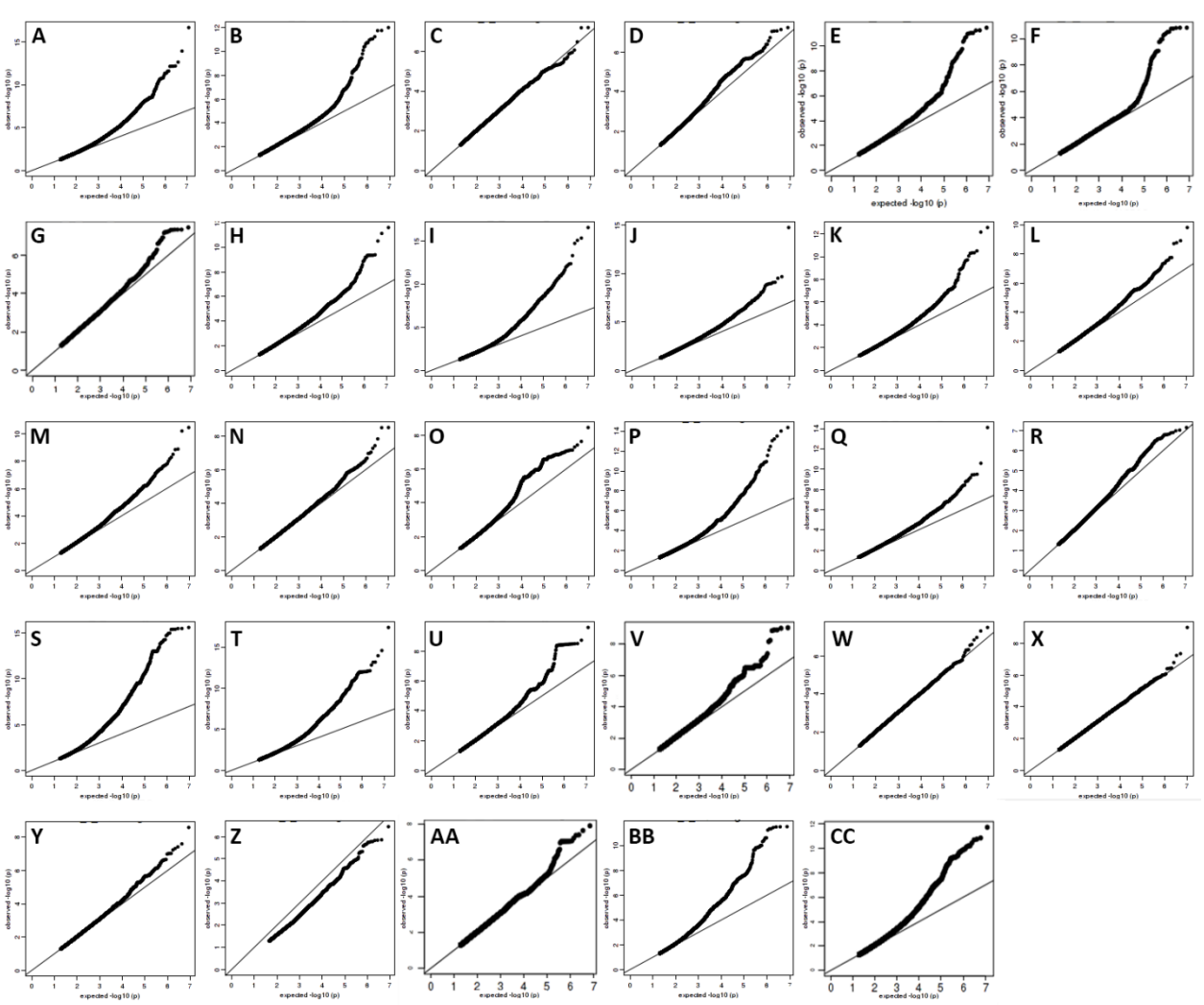
Supplementary Figure 2.1.1: Quality control – Quantile-Quantile (QQ), Pvalue-Z-statistics (P-Z) and Allele Frequency (AF)-plots per cohort in CREAM-ASN generated from EasyQC
Panel 1: QQ-plots per cohort in CREAM-ASN; Panel 2: P-Z plots per cohort in CREAM-ASN; Panel 3: AF plots per cohort in CREAM-ASN; A=BES (n=590), B=Nagahama (n=2730); C=SCES (n=1724), D=SIMES (n=2275), E=SINDI (n=2110), F=STARS (n=817), G=SP2-610 (n=871), H=SP2-1M (n=818).



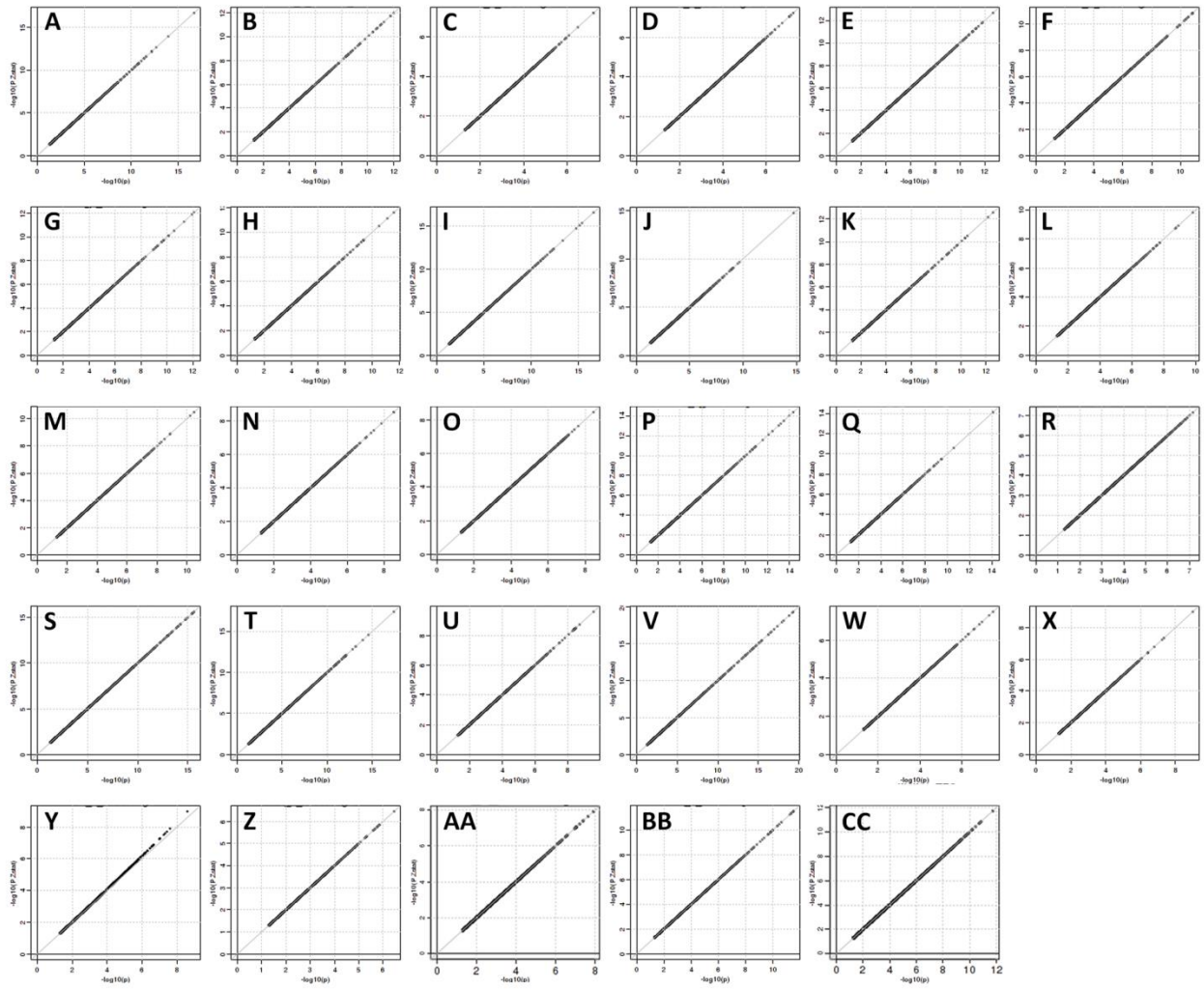
Supplementary Figure 2.1.2: Lambda-N and SE-N plots per cohort in CREAM-ASN generated from EasyQC

A=BES (n=590), B=Nagahama (n=2730); C=SCES (n=1724), D=SIMES (n=2275), E=SINDI (n=2110), F=STARS (n=817), G=SP2-610 (n=871), H=SP2-1M (n=818).

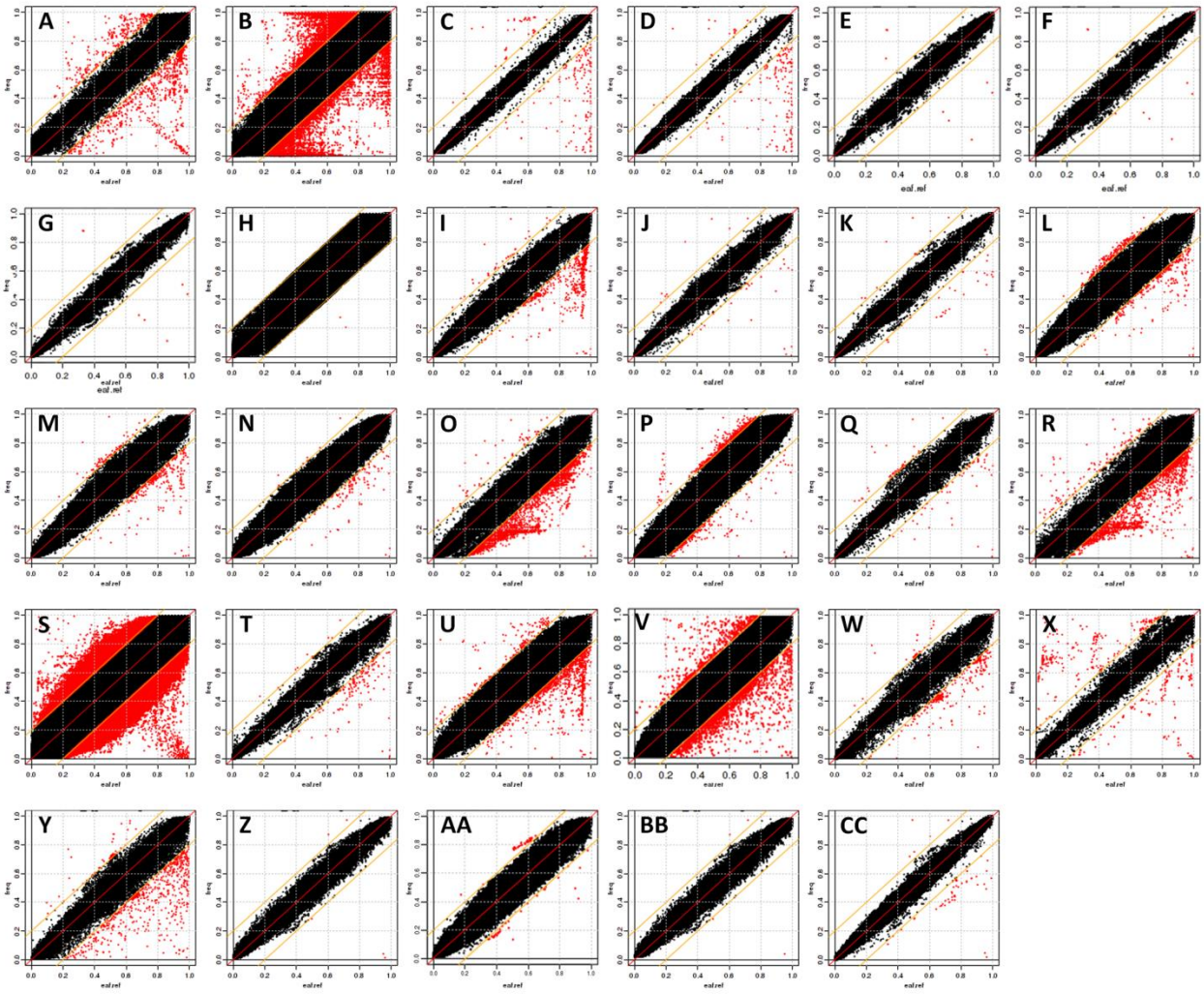
Supplementary Figures 2.2: EasyQC plots per cohort – CREAM-EUR



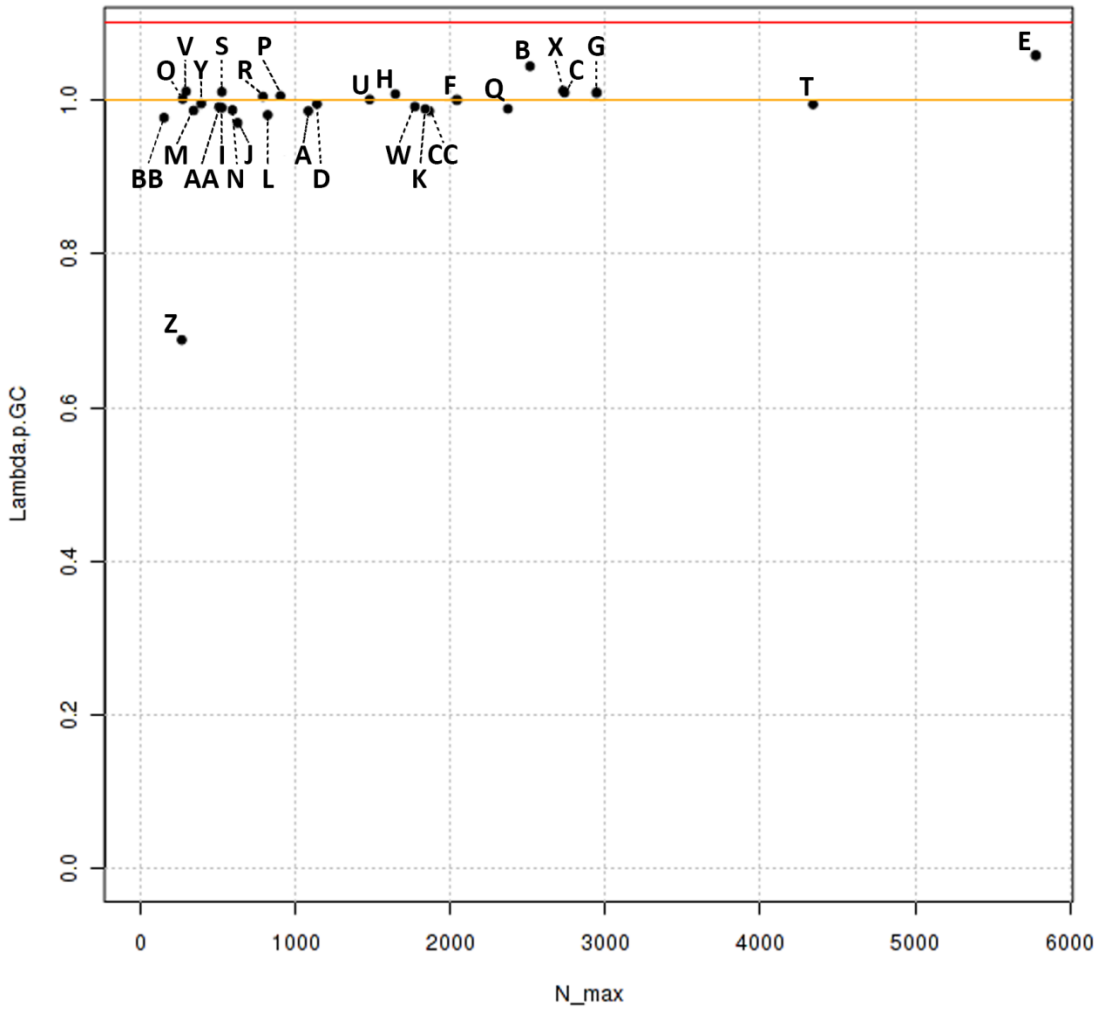
Supplementary Figure 2.2.1: QQ-plots per cohort in CREAM-EUR generated from EasyQC
A=EPIC (n=1084), B=ERF (n=2610), C=GHS1 (n=2738), D=GHS2 (n=1140), E=RSI (n=5787), F=RSII (n=2038), G=RSIII (n=2950), H=1958BBC (n=1658), I=ALIENOR (n=509), J=ANZRAG (n=648), K=ALSPAC (1865), L=CROATIA-KORCULA (n=822), M=CROATIA-SPLIT (n = 344), N=CROATIA-VIS (n=527), O=WESDR (n=295), P=EGCUT (n=904), Q=KORAF (n=2372), R=DCCT (n=791), S=OGP (n=509), T=TWINSUK (n=4342), U=YFS (n=1480), V=FITSA (n=329), W=AREDS (n=1842), X=FRAM (n=2729), Y=FECD (n=393), Z=TEST (n=267), AA=ORCADES (n=1165), BB=BATS (n=158), CC=BMES (n=1896).



Supplementary Figure 2.2.2: PZ-plots per cohort in CREAM-EUR generated from EasyQC
A=EPIC (n=1084), B=ERF (n=2610), C=GHS1 (n=2738), D=GHS2 (n=1140), E=RSI (n=5787), F=RSII (n=2038), G=RSIII (n=2950), H=1958BBC (n=1658), I=ALIENOR (n=509), J=ANZRAG (n=648), K=ALSPAC (1865), L=CROATIA-KORCULA (n=822), M=CROATIA-SPLIT (n = 344), N=CROATIA-VIS (n=527), O=WESDR (n=295), P=EGCUT (n=904), Q=KORAF (n=2372), R=DCCT (n=791), S=OGP (n=509), T=TWINSUK (n=4342), U=YFS (n=1480), V=FITSA (n=329), W=AREDS (n=1842), X=FRAM (n=2729), Y=FECD (n=393), Z=TEST (n=267), AA=ORCADES (n=1165), BB=BATS (n=158), CC=BMES (n=1896).

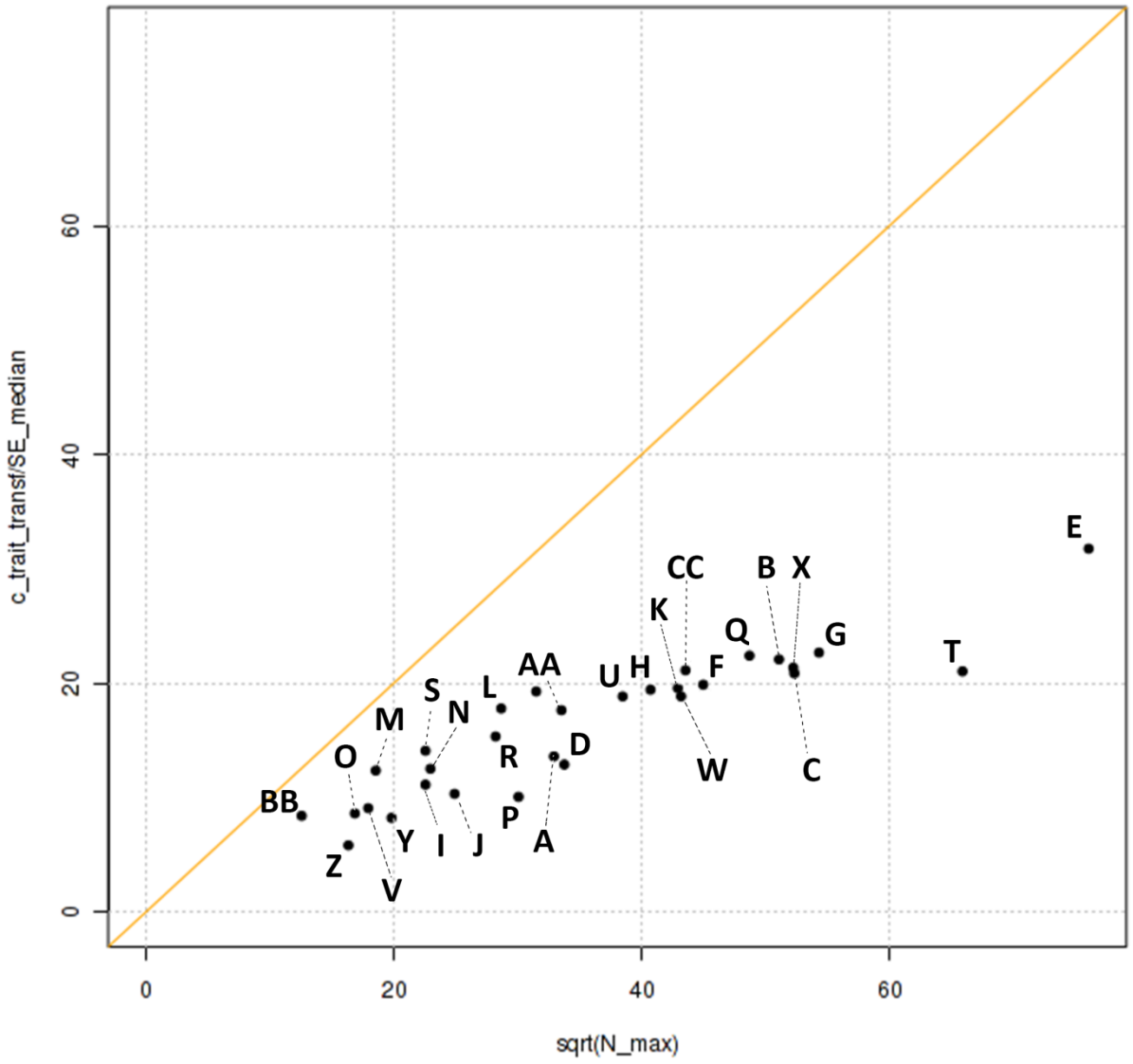


Supplementary Figure 2.2.3: AF-plots per cohort in CREAM-EUR generated from EasyQC
A=EPIC (n=1084), B=ERF (n=2610), C=GHS1 (n=2738), D=GHS2 (n=1140), E=RSI (n=5787), F=RSII (n=2038), G=RSIII (n=2950), H=1958BBC (n=1658), I=ALIENOR (n=509), J=ANZRAG (n=648), K=ALSPAC (1865), L=CROATIA-KORCULA (n=822), M=CROATIA-SPLIT (n = 344), N=CROATIA-VIS (n=527), O=WESDR (n=295), P=EGCUT (n=904), Q=KORAF (n=2372), R=DCCT (n=791), S=OGP (n=509), T=TWINSUK (n=4342), U=YFS (n=1480), V=FITSA (n=329), W=AREDS (n=1842), X=FRAM (n=2729), Y=FECD (n=393), Z=TEST (n=267), AA=ORCADES (n=1165), BB=BATS (n=158), CC=BMES (n=1896).



Supplementary Figure 2.2.4: Lambda-N plot of cohorts in CREAM-EUR generated from EasyQC

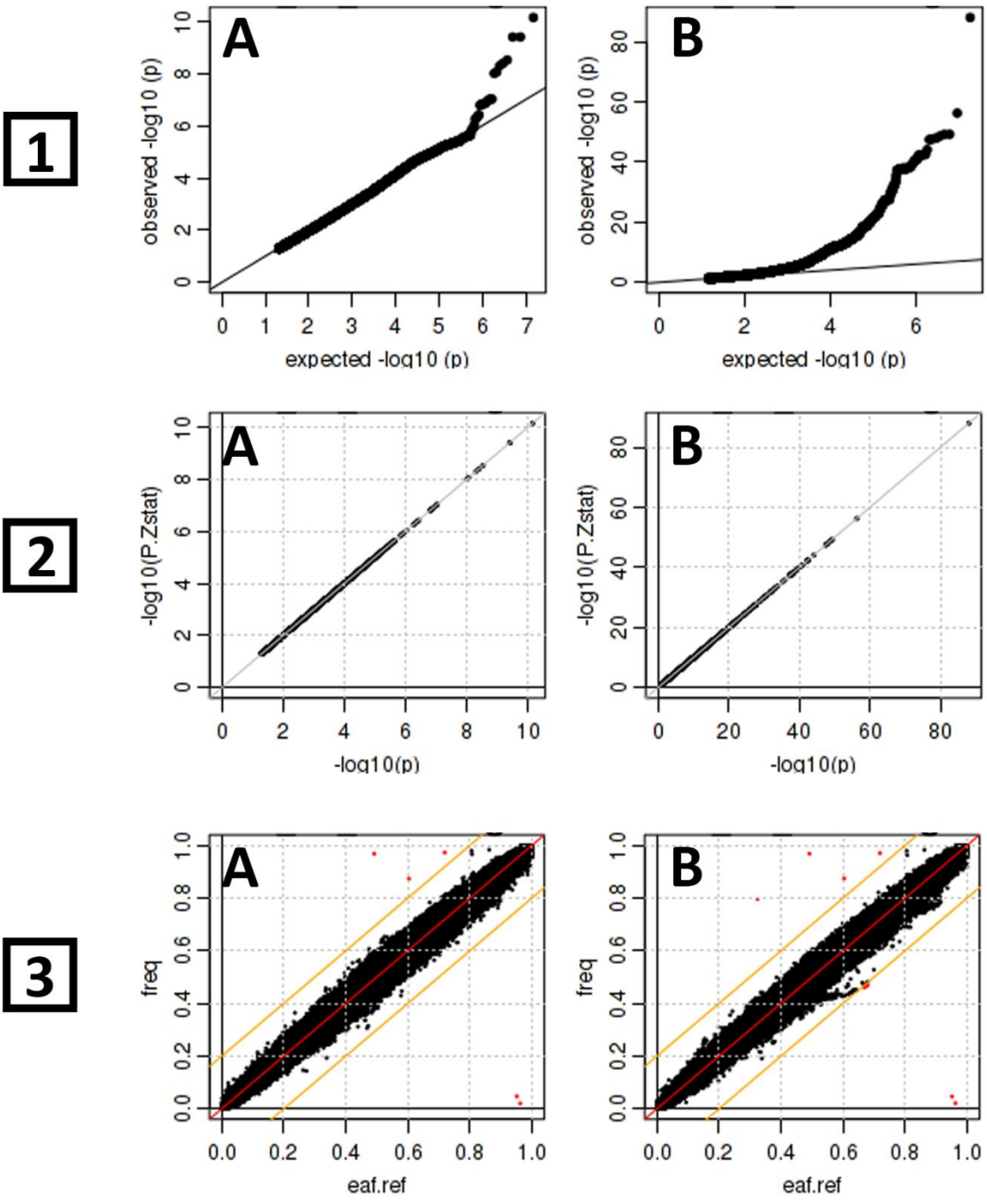
A=EPIC (n=1084), B=ERF (n=2610), C=GHS1 (n=2738), D=GHS2 (n=1140), E=RSI (n=5787), F=RSII (n=2038), G=RSIII (n=2950), H=1958BBC (n=1658), I=ALIENOR (n=509), J=ANZRAG (n=648), K=ALSPAC (1865), L=CROATIA-KORCULA (n=822), M=CROATIA-SPLIT (n = 344), N=CROATIA-VIS (n=527), O=WESDR (n=295), P=EGCUT (n=904), Q=KORAF (n=2372), R=DCCT (n=791), S=OGP (n=509), T=TWINSUK (n=4342), U=YFS (n=1480), V=FITSA (n=329), W=AREDS (n=1842), X=FRAM (n=2729), Y=FECD (n=393), Z=TEST (n=267), AA=ORCADES (n=1165), BB=BATS (n=158), CC=BMES (n=1896).



Supplementary Figure 2.2.5: SE-N plot of cohorts in CREAM-EUR generated from EasyQC

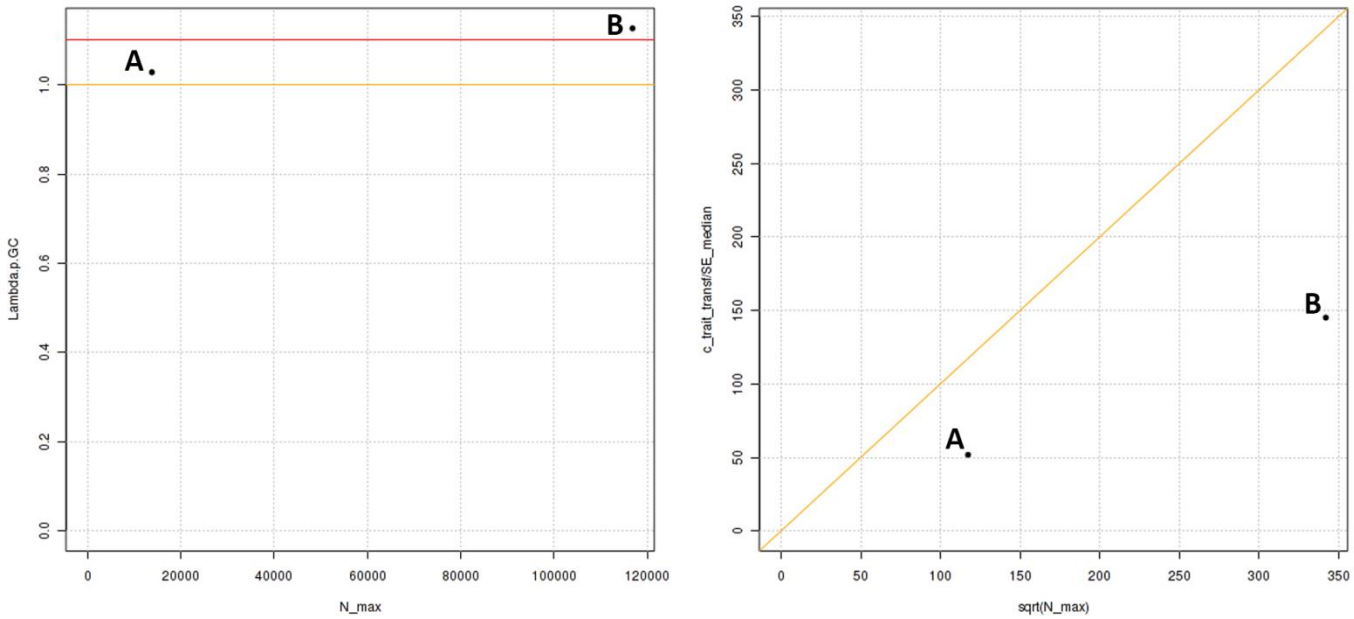
A=EPIC (n=1084), B=ERF (n=2610), C=GHS1 (n=2738), D=GHS2 (n=1140), E=RSI (n=5787), F=RSII (n=2038), G=RSIII (n=2950), H=1958BBC (n=1658), I=ALIENOR (n=509), J=ANZRAG (n=648), K=ALSPAC (1865), L=CROATIA-KORCULA (n=822), M=CROATIA-SPLIT (n = 344), N=CROATIA-VIS (n=527), O=WESDR (n=295), P=EGCUT (n=904), Q=KORAF (n=2372), R=DCCT (n=791), S=OGP (n=509), T=TWINSUK (n=4342), U=YFS (n=1480), V=FITSA (n=329), W=AREDS (n=1842), X=FRAM (n=2729), Y=FECD (n=393), Z=TEST (n=267), AA=ORCADES (n=1165), BB=BATS (n=158), CC=BMES (n=1896).

Supplementary Figure 2.3: EasyQC plots per cohort – 23andMe



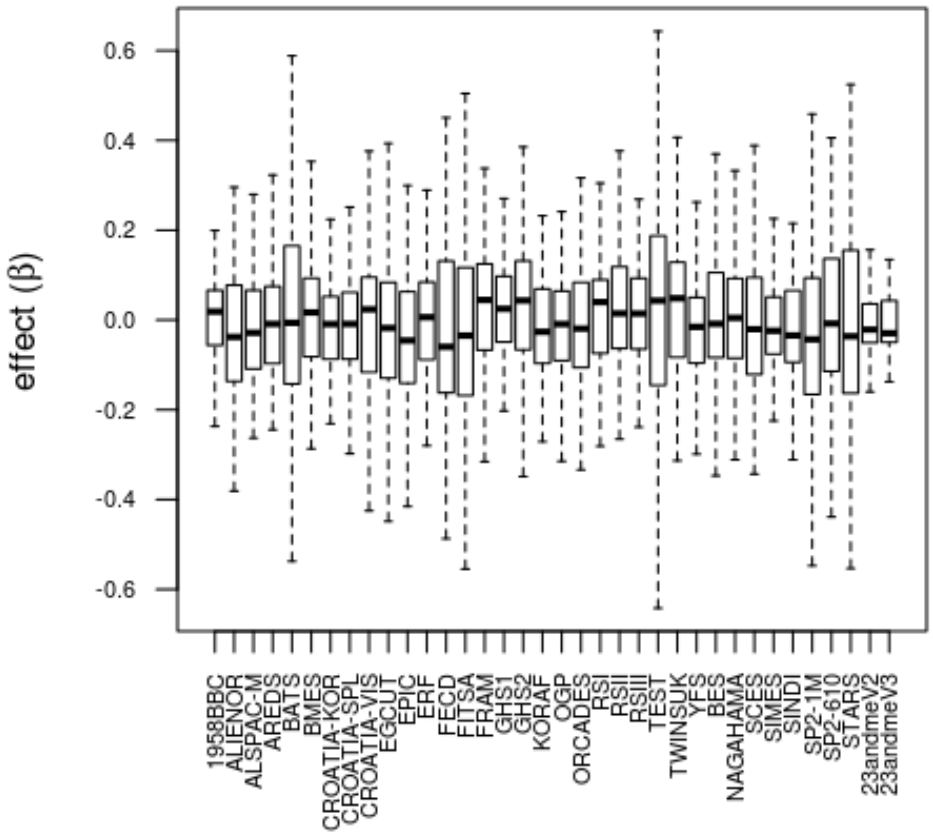
Supplementary Figure 2.3.1: Quality control – QQ, P-Z, AF-plots per cohort in 23andMe generated from EasyQC

Panel 1: QQ-plots per cohort in 23andMe; Panel 2: P-Z plots per cohort in 23andMe; Panel 3: AF plots per cohort in 23andMe; A=23andMe_V2 (n=12128), B=23andMe_V3 (n=92165).



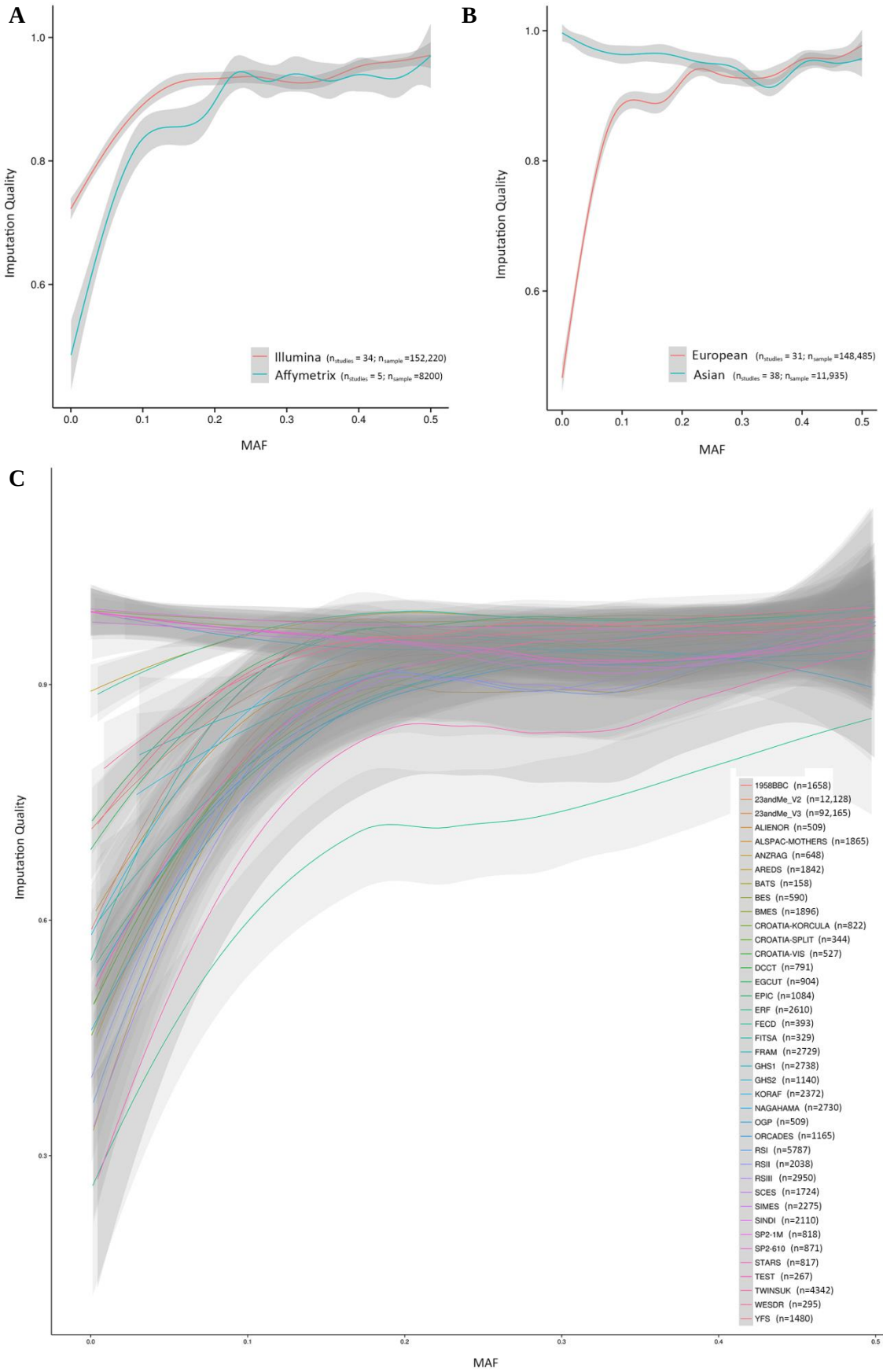
Supplementary Figure 2.3.2: Lambda-N and SE-N plots per cohort in 23andMe generated from EasyQC
A=23andMe_V2 (n=12128), B=23andMe_V3 (n=92165).

Supplementary Figure 3: Boxplots of effect sizes per cohort

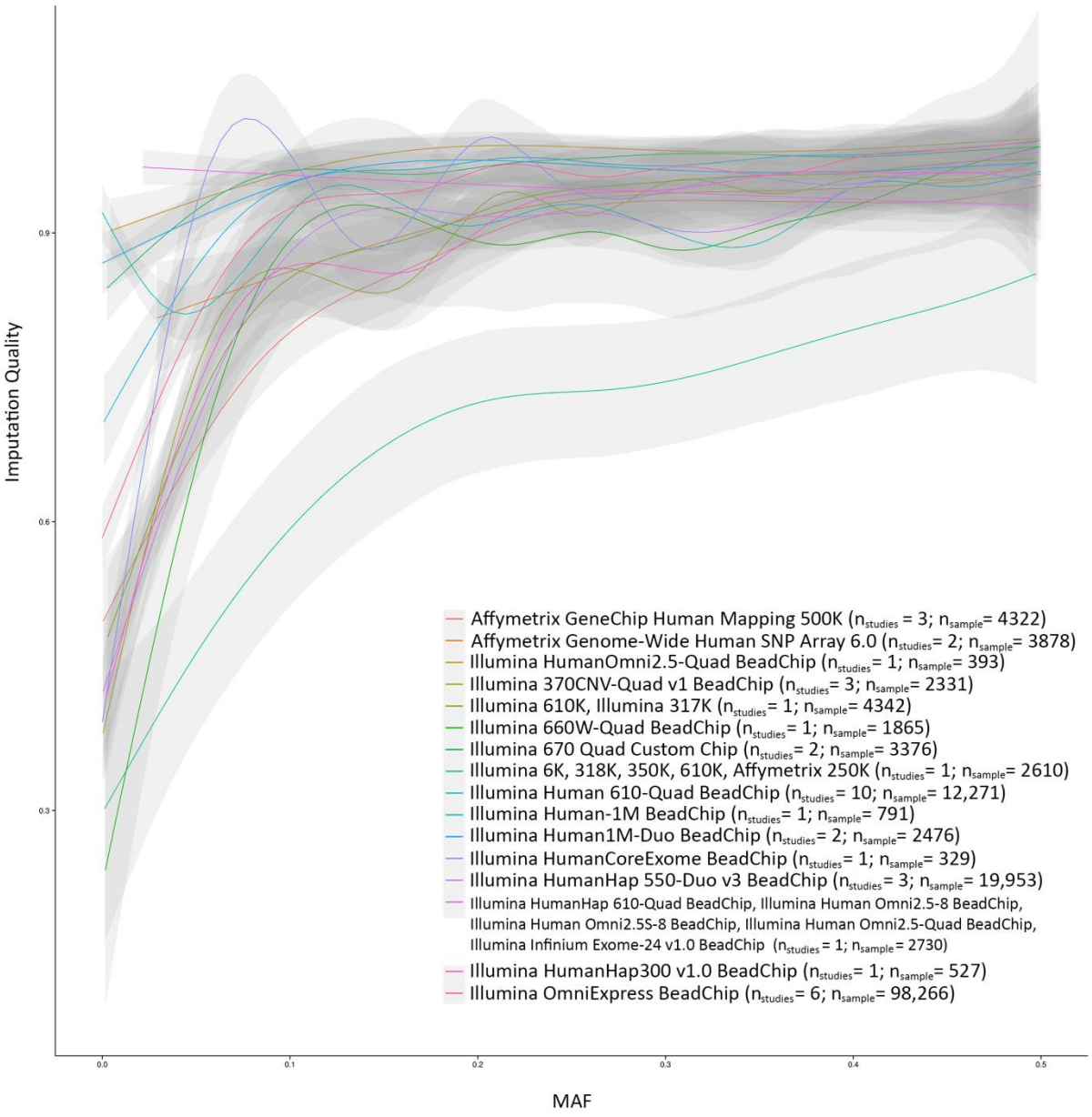


Tukey style box plots of the effect sizes of the 167 independent genetic variants associated with refractive error and myopia derived from the meta-analysis of stage 3 (n=160,420). The bottom and top of the box depict the first and third quartiles, the band is the median and the whiskers extend to the lowest and highest data points. Outliers were not plotted.

Supplementary Figure 4: Imputation Quality of genetic variants of Stage 3



D

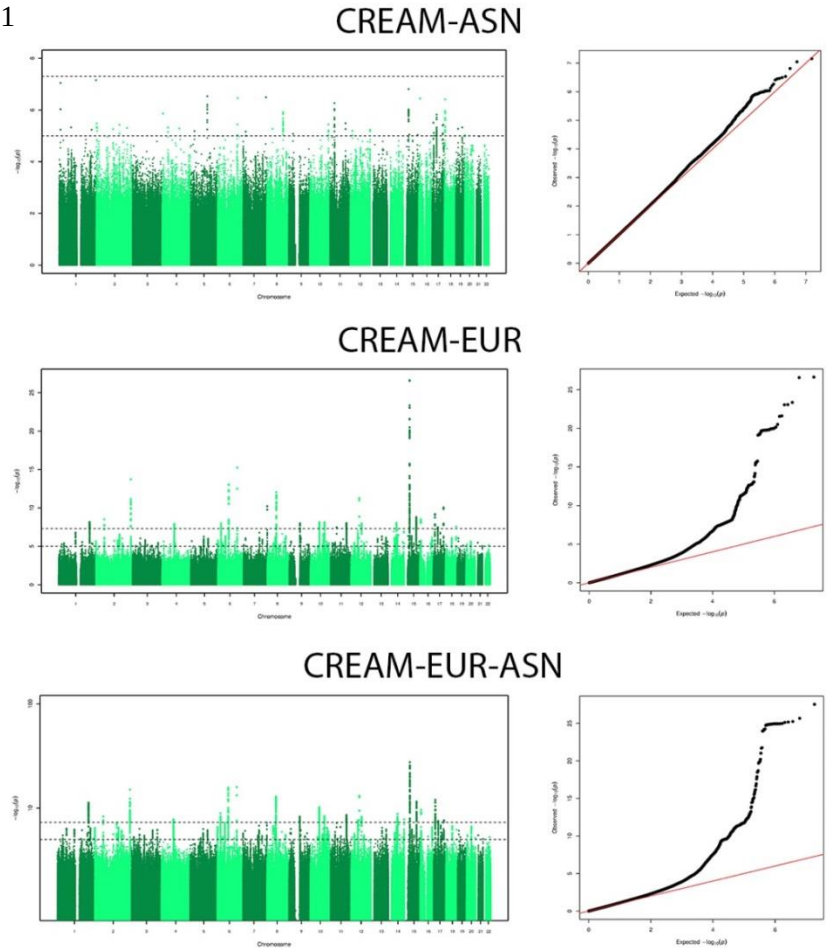


Imputation quality per MAF computed in R (ggplot2)

A) Overall imputation quality (r^2) of the 167 genetic variants before filtering (i.e. $r^2 > 0.3$ and MAF < 0.01 CREAM or < 0.001 23andMe) of the stage 3 meta-analysis of all cohorts using an Illumina or Affymetrix platform, plotted against the minor allele frequency (MAF). **B)** r^2 of the 167 genetic variants before filtering of the stage 3 meta-analysis comparing all European cohorts and Asian, plotted against the minor allele frequency (MAF). **C)** r^2 of the 167 genetic variants before filtering of the stage 3 meta-analysis of all participating cohorts. **D)** r^2 of the 167 genetic variants before filtering of the stage 3 meta-analysis based on type of array used. Bands indicate the 95% confidence intervals; $n_{studies}$: number of cohorts; n_{sample}: number of participants tested.

Supplementary Figure 5: Manhattan plots and QQ plots per Stage (1-3)

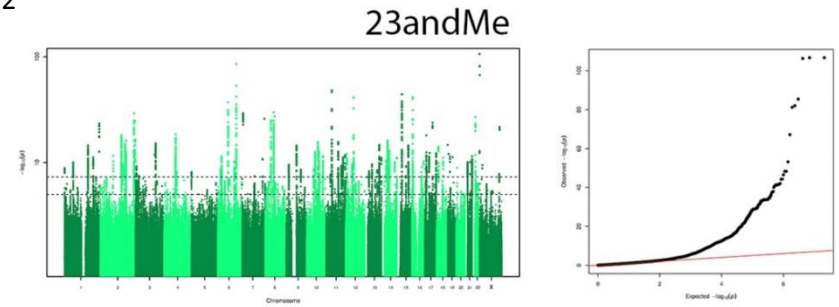
Stage 1



Shown are the Manhattan plots and QQ plots depicting P values for association of the meta-analysis of Stage 1.

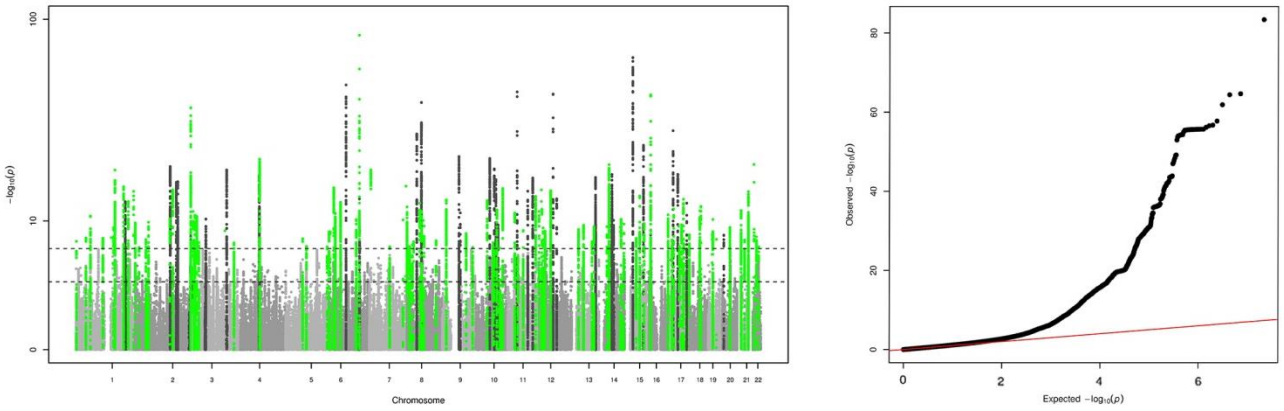
CREAM-EUR (n_{sample}= 44,192 participants; $n_{genetic\ variants}$ =9.6M); $\lambda_{genomic\ inflation}$ = 1.119. CREAM-ASN (n_{sample} = 11,935 participants; $n_{genetic\ variants}$ = 7.8M); $\lambda_{genomic\ inflation}$ = 1.022. CREAM-EUR-ASN (n_{sample} = 56,127 participants; $n_{genetic\ variants}$ = 9.3M); $\lambda_{genomic\ inflation}$ = 1.124. The red lines in the QQ plots indicate x=y.

Stage 2



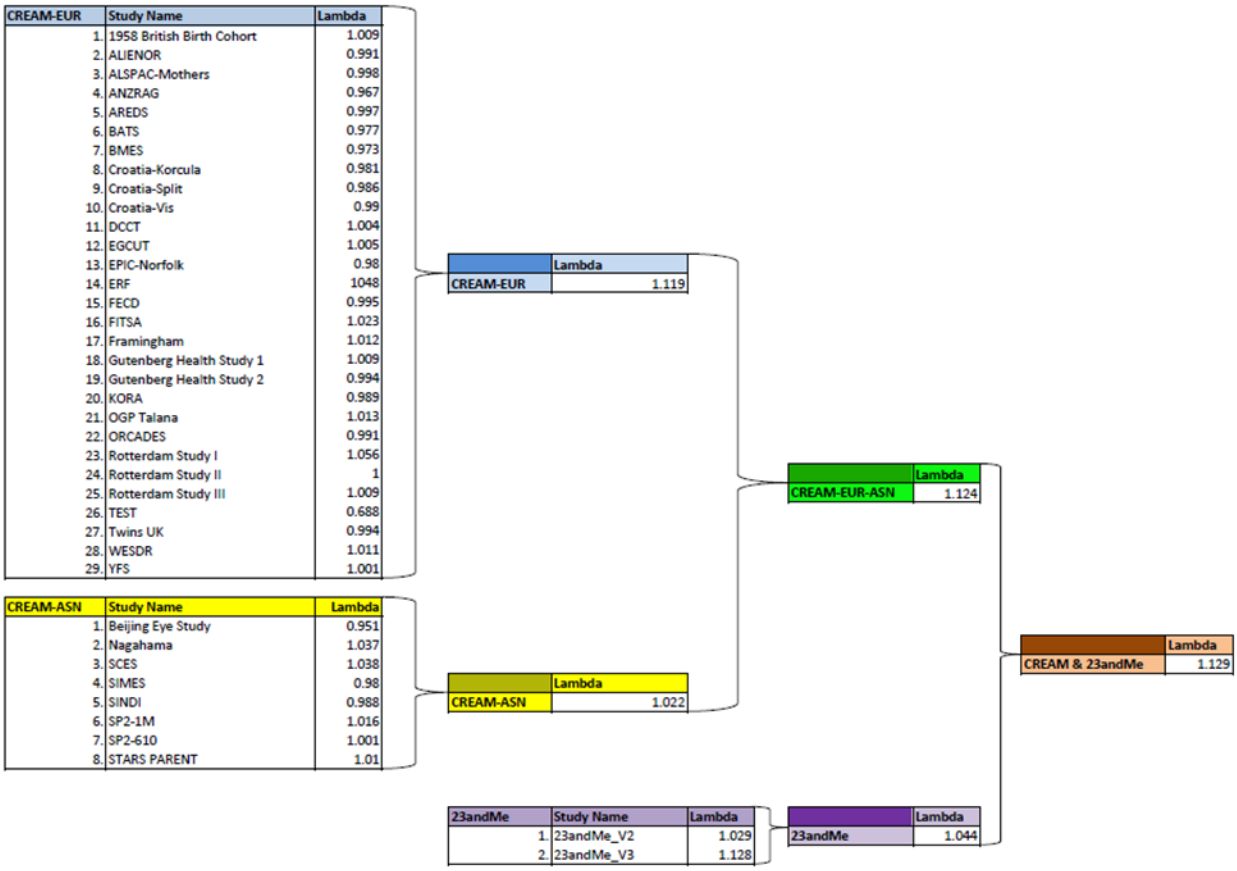
23andMe (n_{sample}= 104,293 participants; $n_{genetic\ variants}$ = 10.6M); $\lambda_{genomic\ inflation}$ = 1.119. The red line in the QQ plot indicates x=y.

Stage 3

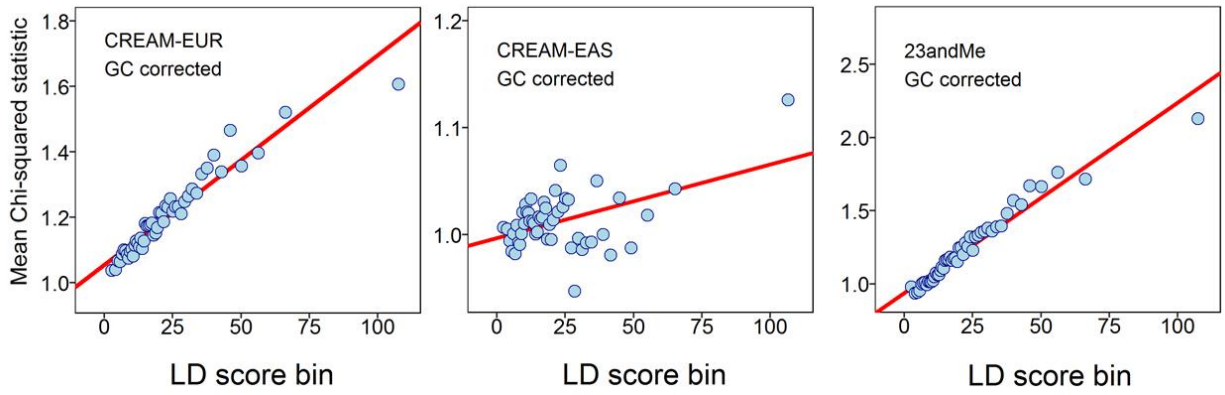


Shown is the Manhattan plot depicting P values for association of the meta-analysis of CREAM-EUR&ASN and 23andMe (n_{sample}=160,420 participants; $n_{genetic\ variants}$ = 11M), highlighting new ($P < 5 \times 10{-}8$ for the first time; green) and known (dark grey) refractive error loci previously; $\lambda_{genomic\ inflation}$ = 1.129. The red line indicates x=y.

Supplementary Figure 6: Lambdas of all cohorts and fixed effects meta-analyses

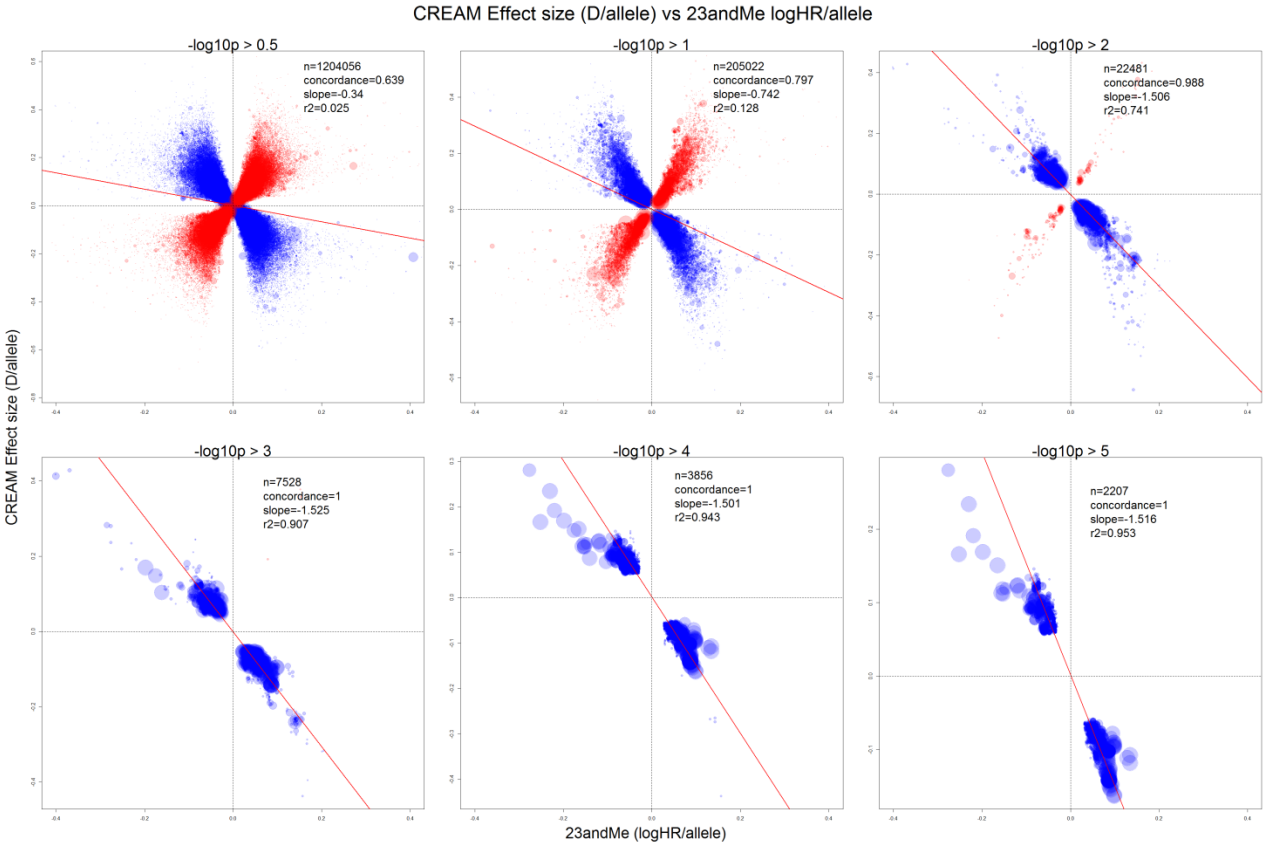


Supplementary Figure 7: LD-score regressions per ancestry of CREAM and 23andMe



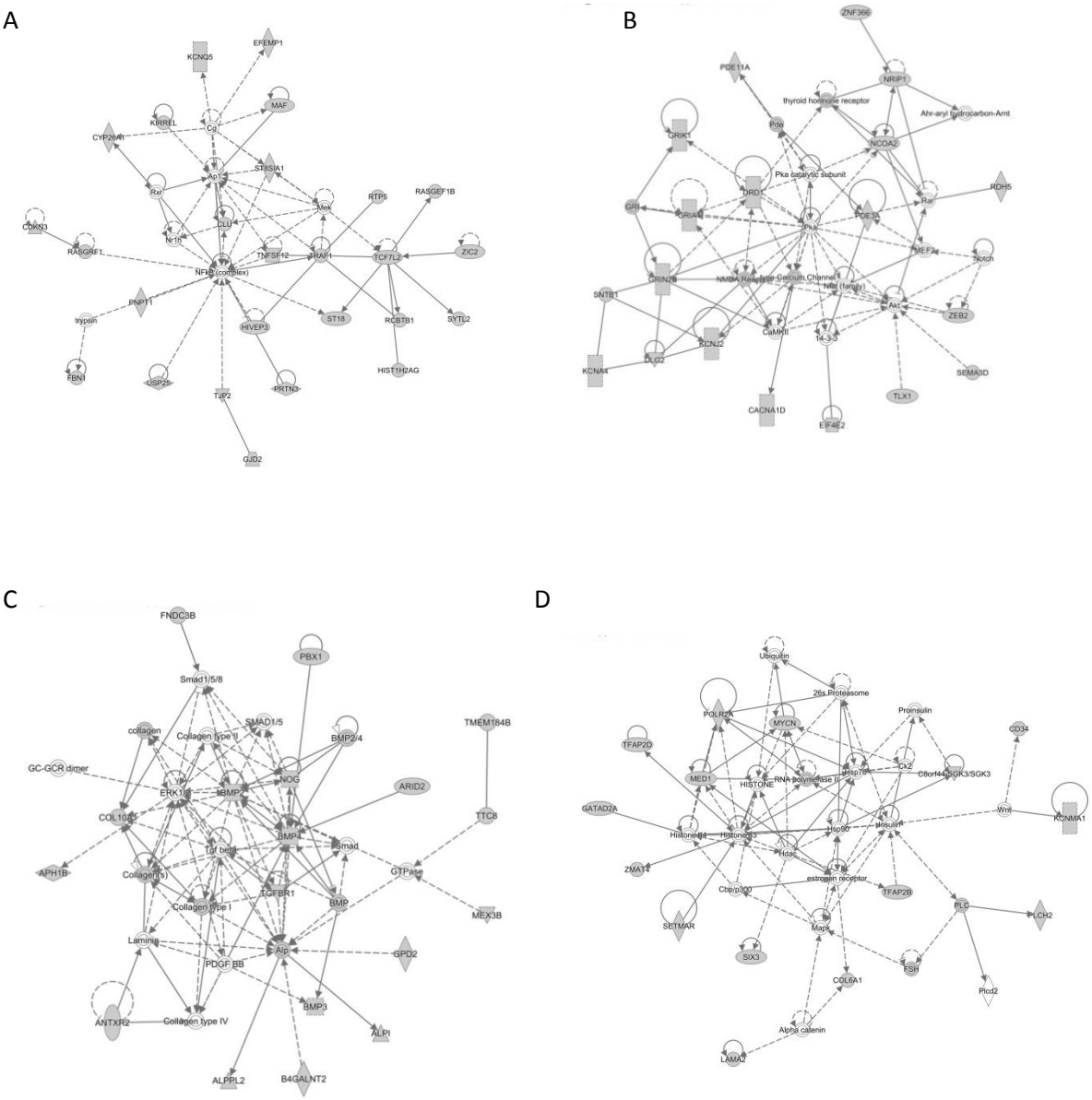
We used Popcorn[57] to investigate ancestry-related differences in the genetic architecture of refractive error and myopia. Pairwise analyses were carried out using the GWAS summary statistics from the 23andMe (n=104,292), CREAM-EUR (n=44,192) and CREAM-EAS (n=9,826) meta-analyses. Only SNPs with MAF ≥ 5% were included, resulting in a final set of 3,625,602 SNPs for analyses involving 23andMe and 3,642,928 SNPs for the CREAM-EUR vs. CREAM-EAS analysis.

Supplementary Figure 8: Effects comparison SphE and AODM using different p-value thresholds



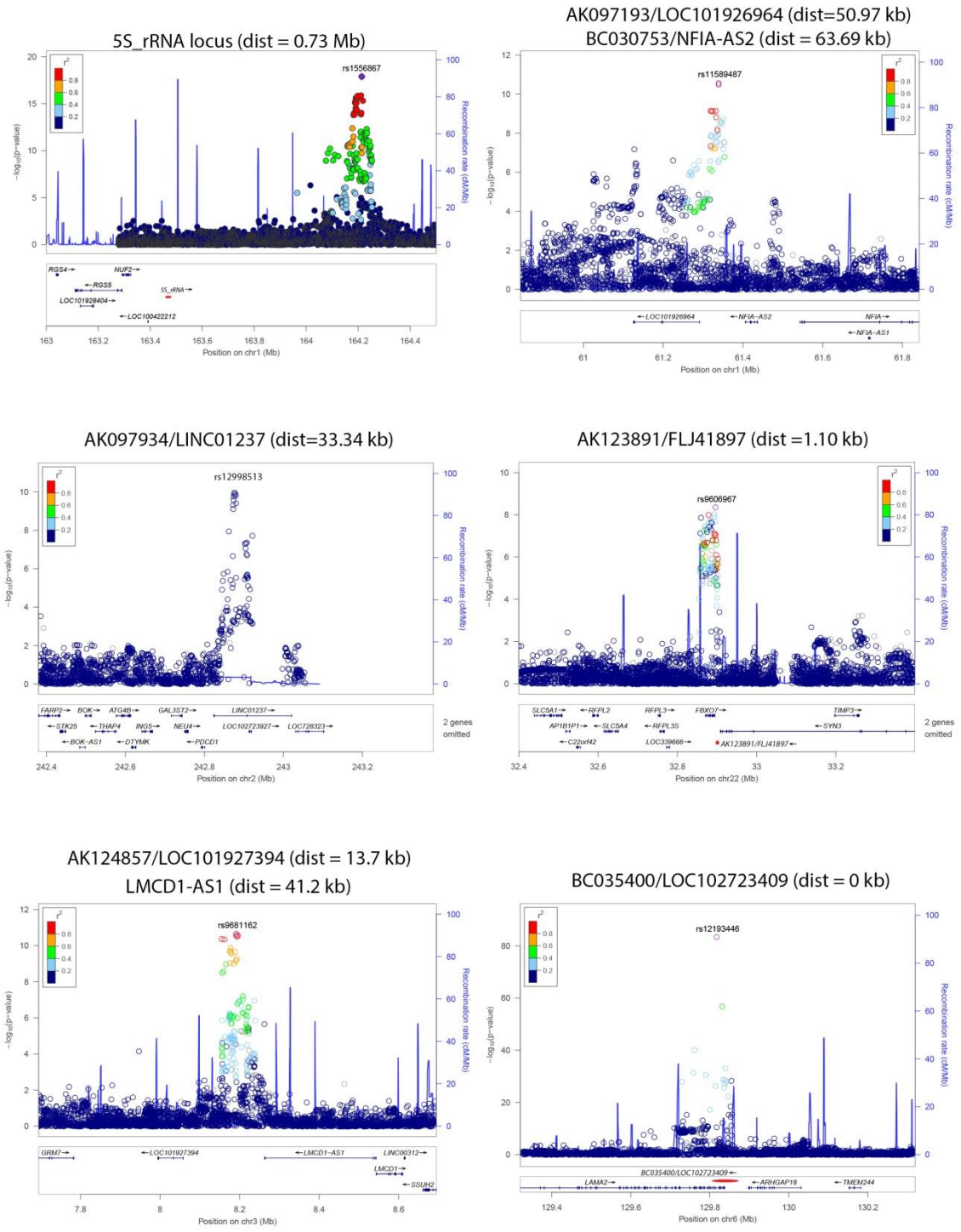
Shown are the graphs for the comparison of the effects, SphE and AODM, using different p-value thresholds. These p-values are derived from the meta-analysis of stage 1 and 2, i.e. CREAM and 23andMe. Red line = regression line; n = number of genetic variants tested at different *P* value thresholds (-log10p > 0.5 – -log10p >5); concordance = concordance of correlation coefficient; slope = slope of regression line; r^2= Pearson correlation coefficient.

Supplementary Figure 9: Top canonical pathways of Ingenuity Pathway Analysis



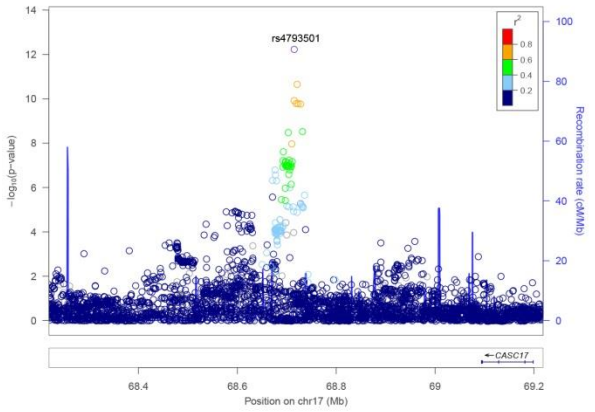
The 197 of the 208 genes annotated to the top hits identified at Stage 3 were present in the IPA database and mapped to networks and pathways. The networks were identified by right-tailed Fisher's exact tests. Genes within the network indicated in grey are genes associated with refractive error. A) Glutamate receptor signaling, *P* value = $1.56E^{-4}$; B) Factors Promoting Cardiogenesis in Vertebrates, *P* value = $1.78E^{-4}$; C) CREB Signaling in Neurons; *P* value = $1.37E^{-3}$; D) Human Embryonic Stem Cell Pluripotency, *P* value = $1.83E^{-3}$

Supplementary Figure 10: Locus Zoom plots of RNA gene regions

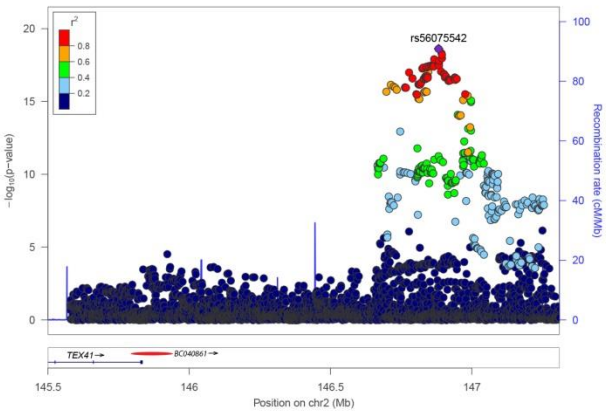


Regional plots of the RNA genes depicting the p-values derived from the meta-analysis of stage 3 (n=160,420), linkage disequilibrium and recombination ratio. Some of the RNA genes were not depicted by the Locus Zoom software, in which case they were added as the red regions between the genes annotated by Locus Zoom; dist, distance; $r^2 = r^2$ of linkage disequilibrium.

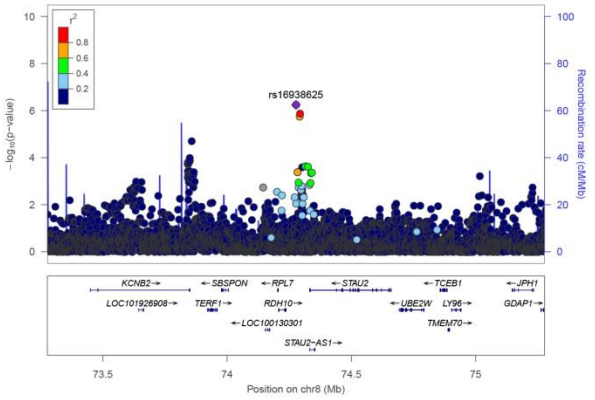
BC039327/CASC17 (dist = 0.38 Mb)



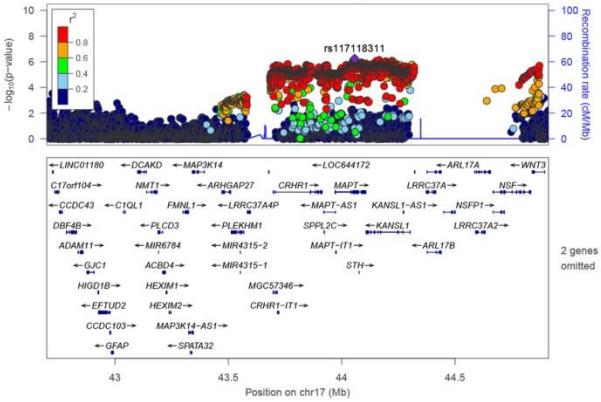
BC040861 (dist = 0.97 Mb)



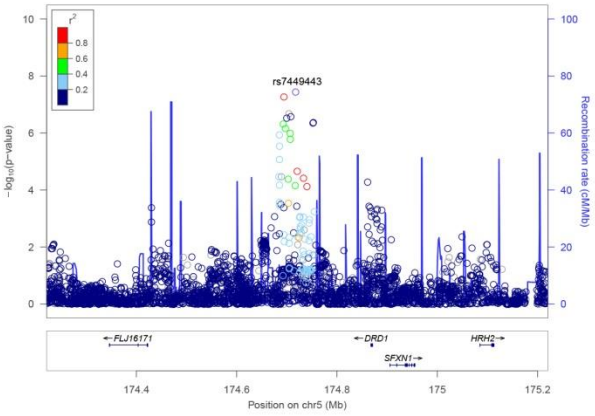
BC127738/LOC100130301 (Gene based finding; dist = 105.1 kb)



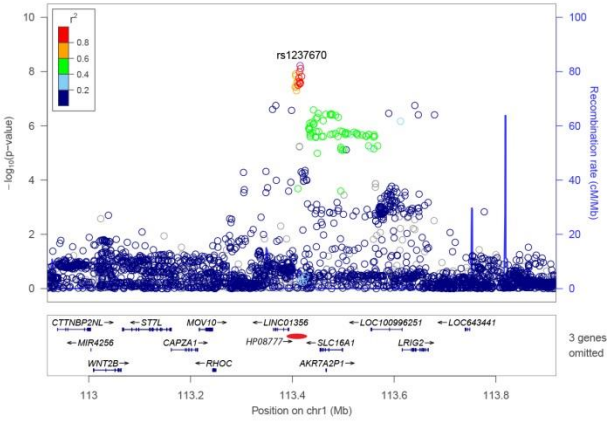
CRHR1-IT1 (Gene based finding; dist = 2.26 Mb)



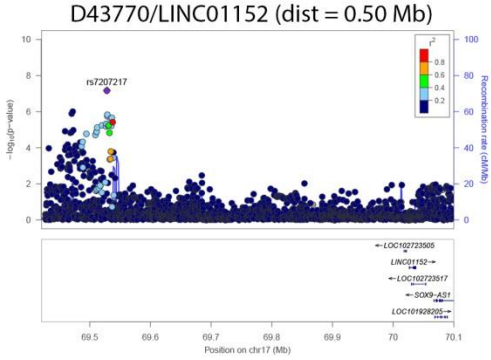
FLJ16171/LINC01951 (dist = 0.30 Mb)



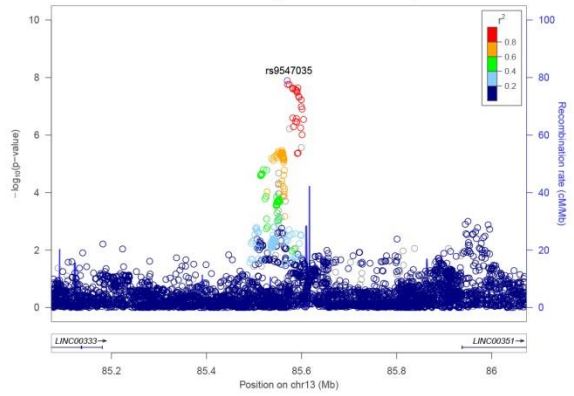
HP08777 (dist = 0 kb)



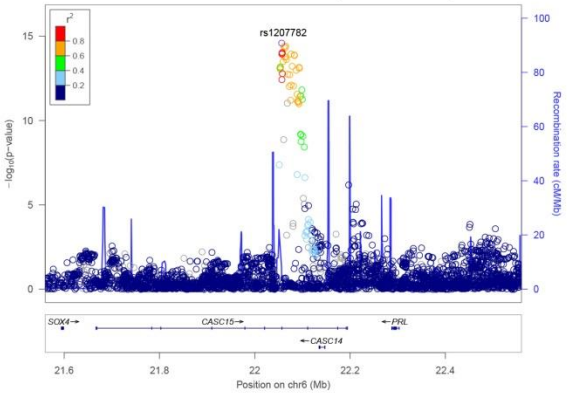
D43770/LINC01152 (dist = 0.50 Mb)



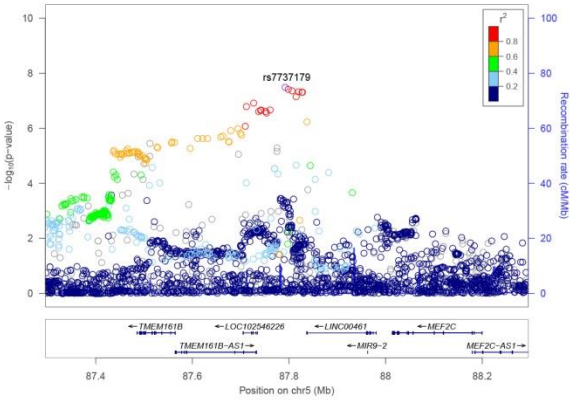
LINC00333 (dist = 393 kb)
LINC00351 (dist = 364 kb)



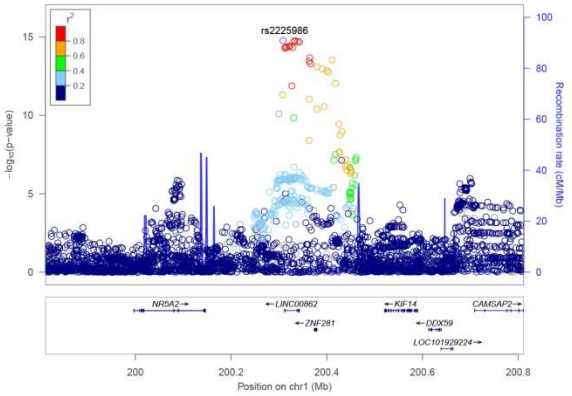
LINC00340/CASC15 (dist = 0 kb)



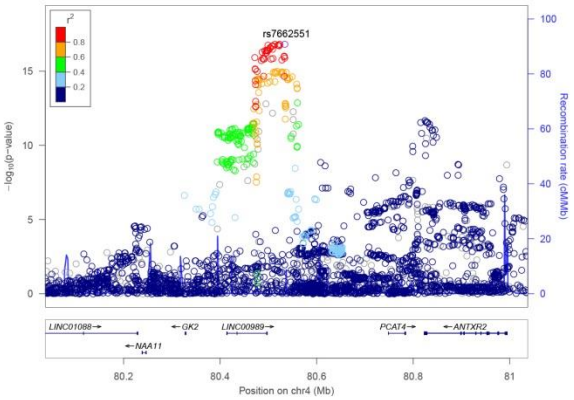
LINC00461 (dist = 41 kb)
TMEM161B-AS1 (dist = 63 kb)



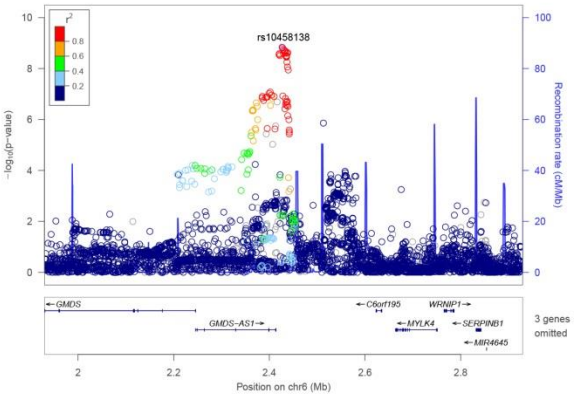
LINC00862 (dist = 0 kb)



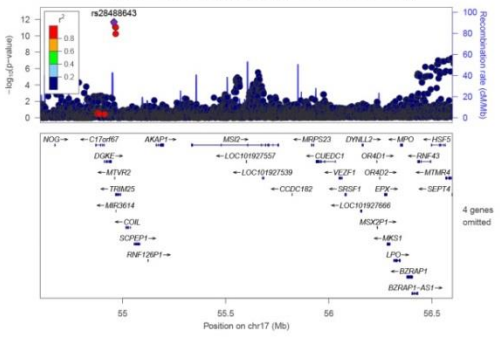
LOC100506035/LINC00989 (dist = 40 kb)



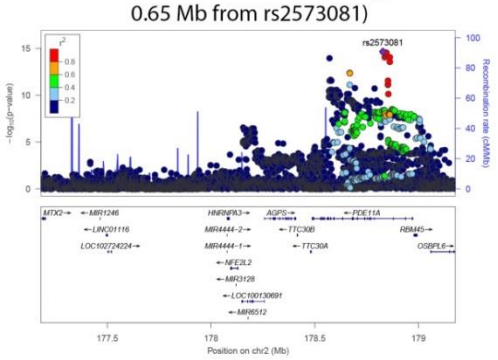
LOC100508120/GMDS-AS1 (dist = 0 kb)



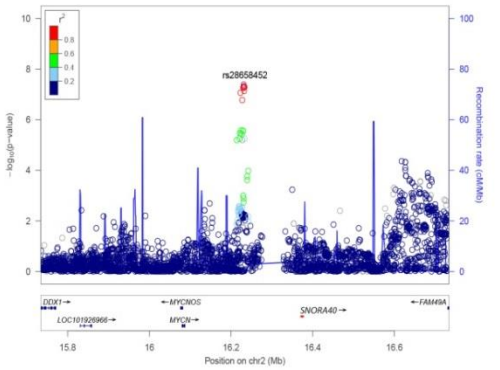
LOC101927557 (Gene based finding; dist = 0.64 Mb from rs28488643)



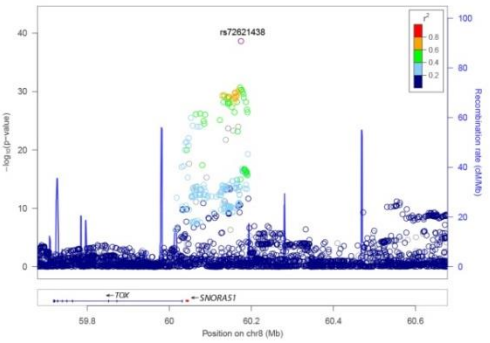
MIR6512 (Gene based finding; 0.65 Mb from rs2573081)



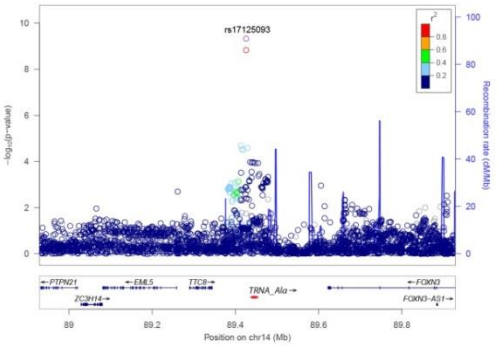
SNORA40 (dist = 146 kb)



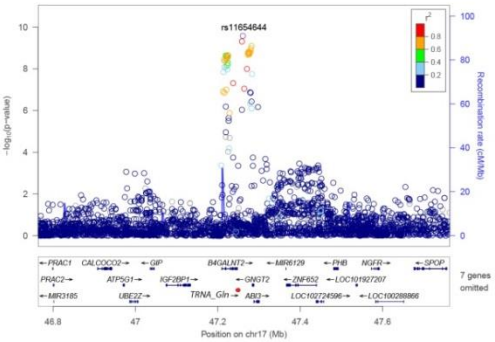
SNORA51 (dist = 129 kb)



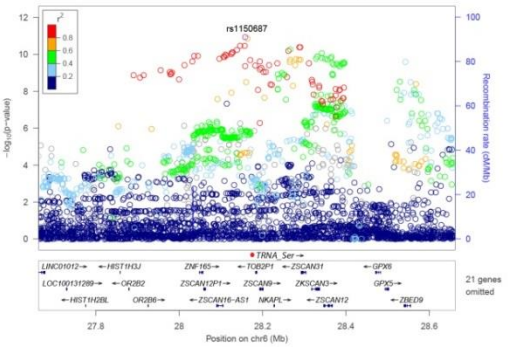
TRNA_Ala (dist = 16.5 kb)



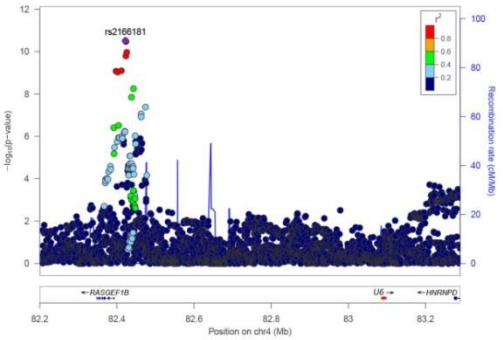
TRNA_Gln (dist = 6.4 kb)



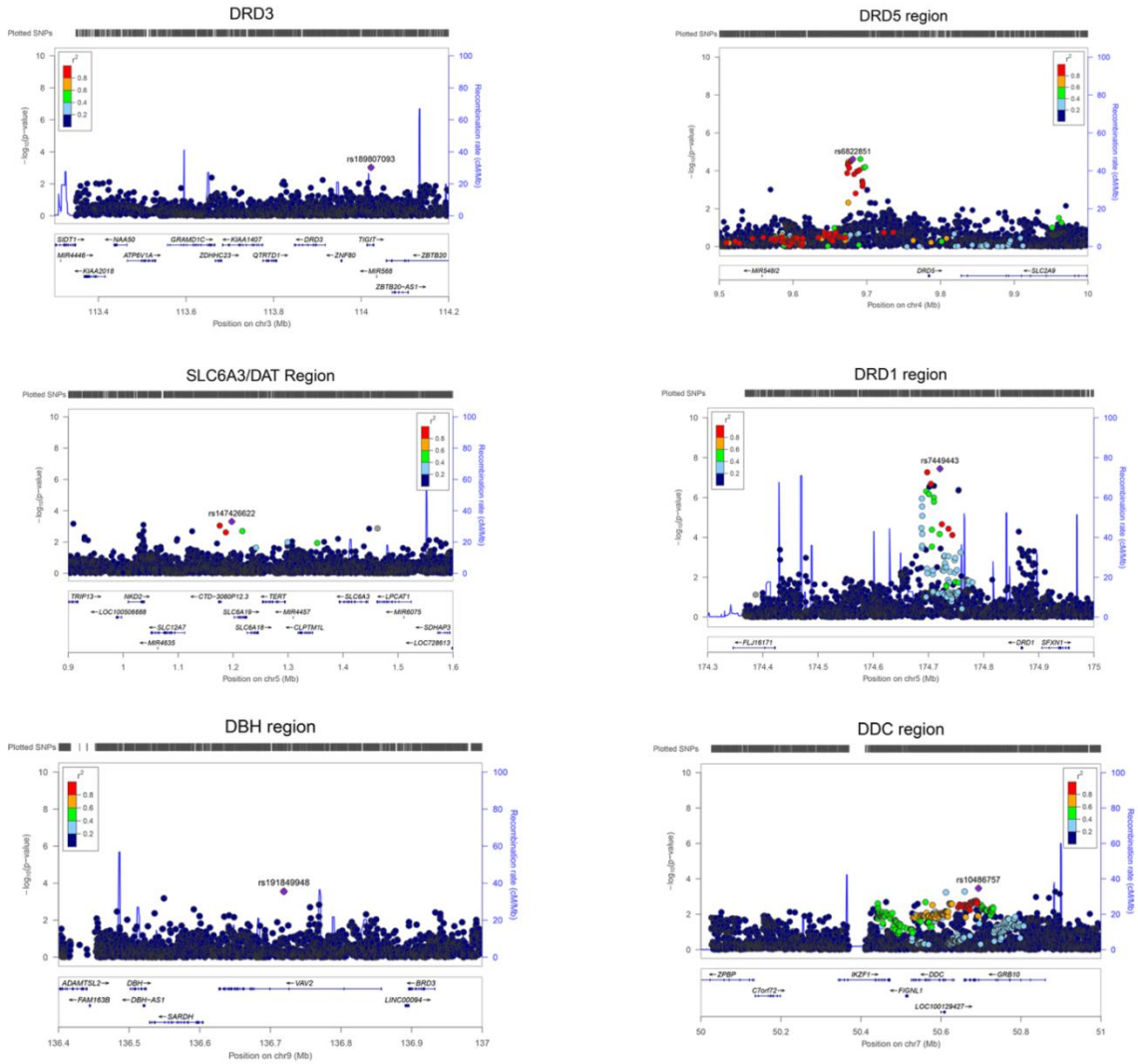
TRNA_Ser (dist = 18.4 kb)



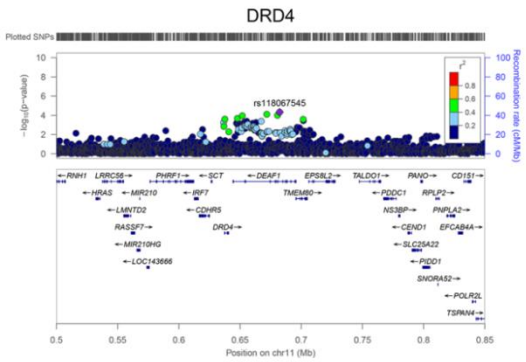
U6 (dist = 0.67 Mb)

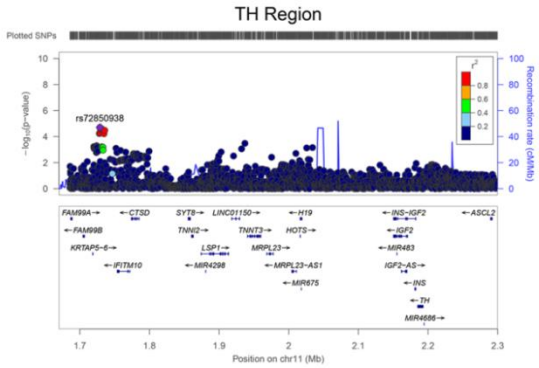


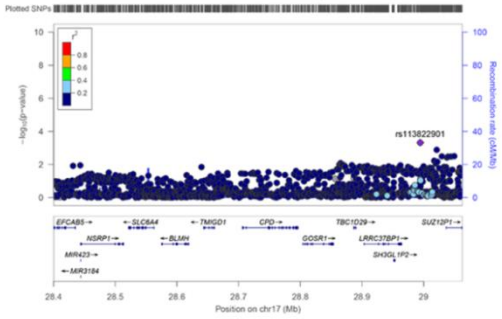
Supplementary Figure 11: Locus Zoom plots of dopamine pathway regions

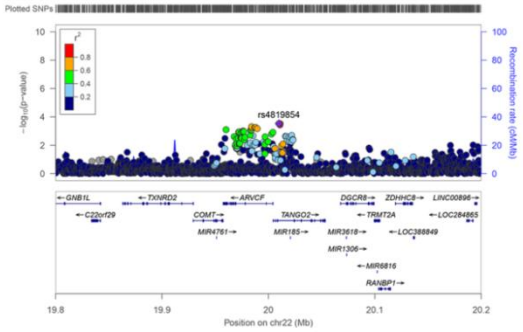


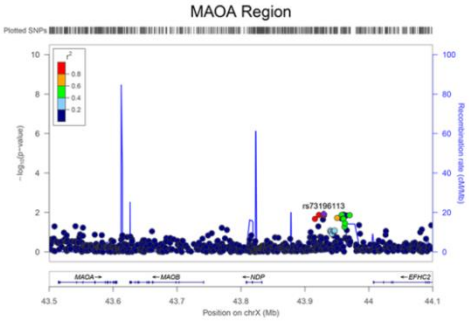
Regional plots of the key genes in the dopamine pathway depicting the p-values derived from the meta-analysis of stage 3 (n=160,420), linkage disequilibrium and recombination ratio; r^2 = r^2 of linkage disequilibrium.


DRD4
Plotted SNPs
rs11067545
Position on chr11 (Mb)

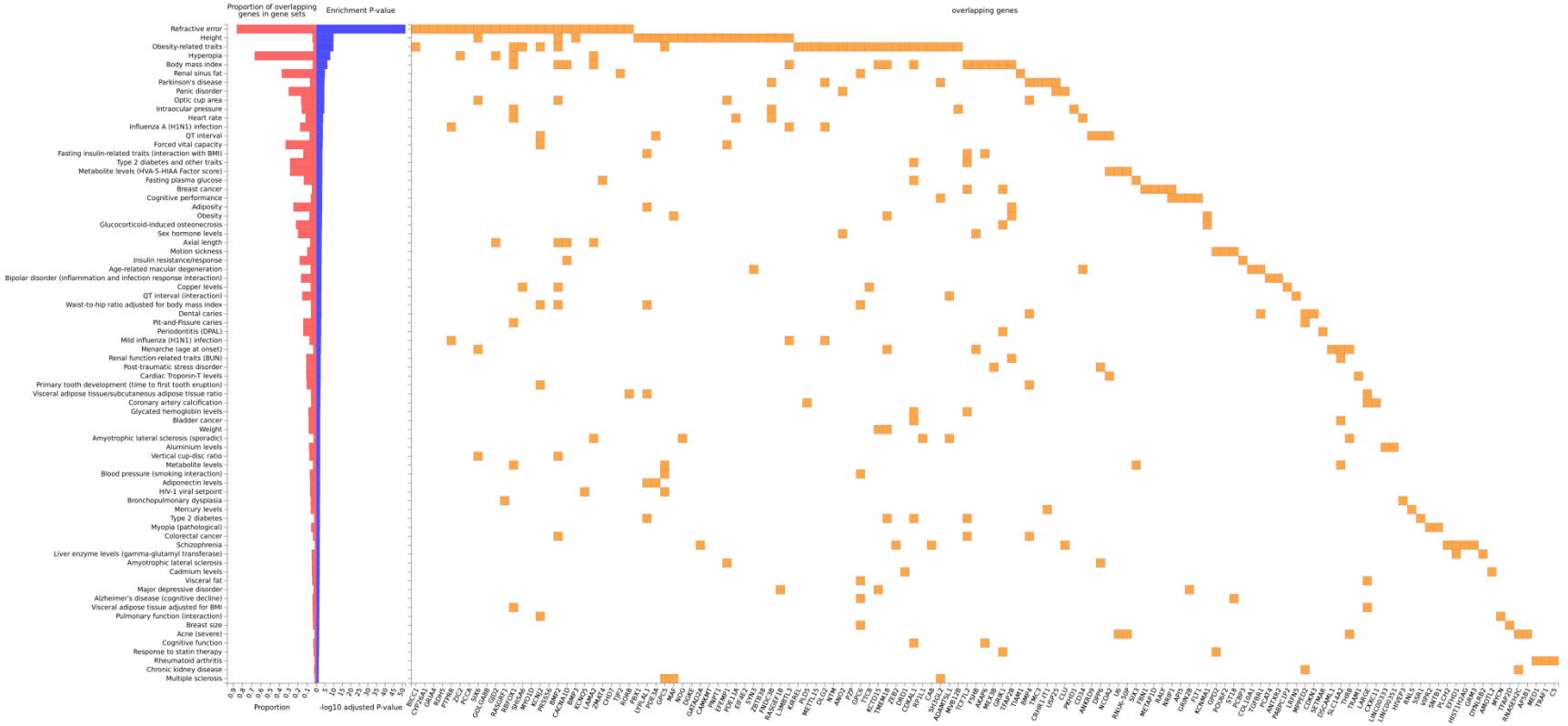

TH Region
Plotted SNPs
rs72850938
Position on chr11 (Mb)


SLC6A4/SERT Region
Plotted SNPs
rs113822901
Position on chr17 (Mb)


COMT region
Plotted SNPs
rs4819854
Position on chr22 (Mb)


MAOA Region
Plotted SNPs
rs73196113
Position on chrX (Mb)

Supplementary Figure 12: GWAS catalog comparison of refractive error genes with other diseases



Supplementary Table 1a: Descriptives per cohort

CREAM-EUR

Study Name	Origin	Study design	n	age (sd)	% female	mean SE
1958 British Birth Cohort	United Kingdom	population based study	1658	42.00 (0.00)	46.0	-0.96 (2.003)
ALIENOR	France	population based study	509	79.15 (4.06)	56.8	0.98 (1.97)
ALSPAC-Mothers	United Kingdom	population based study	1865	45.02 (4.53)	100.0	-0.76 (2.16)
ANZRAG	Australia	POAG cases	648	79.02 (12.08)	49.3	-0.21 (2.41)
AREDS	United States	population based study	1842	68.08 (4.71)	59.0	0.54 (2.16)
BATS	Australia	population based twin study	158	26.52 (2.41	56.3	-0.51 (1.15)
BMES	Australia	population based study	1896	67.09 (9.16)	57.3	0.62 (2.12)
Croatia-Korcula	Croatia	population based study	822	56.33 (13.34)	64.8	-0.15 (1.60)
Croatia-Split	Croatia	population based study	344	51.95 (13.02)	61.1	-1.27 (1.57)
Croatia-Vis	Croatia	population based study	527	56.29 (13.30)	60.0	-0.13 (1.74)
DCCT	United States	clinical trial	791	31.43 (4.13)	43.2	-1.47 (0.80)
EGCUT	Estonia	population based study	904	56.00 (17.00)	61.2	0.33 (3.36)
EPIC-Norfolk	United Kingdom	population based study	1084	68.81 (7.55)	56.3	0.34 (2.27)
ERF	The Netherlands	family based study	2610	48.72 (14.17)	55.0	0.13 (2.03)
FECD	United States	case-control (controls only)	393	71.50 (9.18)	60.2	-0.14 (2.49)
FITSA	Finland	population based twin study	329	68.56 (3.35)	100.0	1.22 (1.71)
Framingham	United States	population based study	2729	55.60 (8.90)	42.5	0.03 (2.410)
Gutenberg Health Study 1	Germany	population based study	2738	55.52 (10.81)	48.6	-0.38 (2.45)
Gutenberg Health Study 2	Germany	population based study	1140	54.81 (10.81)	50.4	-0.41 (2.57)
KORA	Germany	population based study	2372	55.14 (11.79)	67.0	-0.25 (2.22)
OGP Talana	Italy	population based study	509	51.43 (19.51)	59.2	-0.10 (1.67)
ORCADES	United Kingdom	population based study	1165	55.83 (13.76)	61.0	0.09 (2.07)
Rotterdam Study I	The Netherlands	population based study	5787	68.84 (8.84)	59.4	0.83 (2.55)
Rotterdam Study II	The Netherlands	population based study	2038	64.24 (7.75)	54.4	0.49 (2.49)
Rotterdam Study III	The Netherlands	population based study	2950	56.91 (6.54)	55.9	-0.28 (2.60)
TEST	Australia	population based twin study	267	46.10 (12.25)	50.3	-0.54 (1.99)
Twins UK	United Kingdom	population based twin study	4342	53.83 (11.12)	92.2	-0.34 (2.72)
WESDR	United States	case-control from population based study	295	34.63 (8.05)	51.2	-1.53 (2.02)
YFS	Finland	population based study	1480	41.94 (5.02)	55.4	-1.02 (1.99)
			44192			

CREAM-ASN

Study Name	Origin	Study design	n	age (sd)	% female	mean SE
Beijing Eye Study	China	population based study	590	62.13 (8.51)	66.1	-0.06 (1.87)
Nagahama	Japan	population based study	2730	51.29 (14.03)	66.6	-1.69 (2.78)
SCES	Singapore (Chinese)	population based study	1724	57.54 (8.99)	48.6	-0.77 (2.65)
SIMES	Singapore (Malay)	population based study	2275	58 (10.81)	50.9	-0.05 (1.86)
SINDI	Singapore (Indian)	population based twin study	2110	55.82 (8.82)	48.6	0.01 (2.14)
SP2-1M	Singapore	population based study	818	46.81 (10.16)	37.7	-1.81 (2.85)
SP2-610	Singapore	population based study	871	48.54 (11.32)	80.4	-1.51 (3.00)
STARS PARENT	Singapore	population based study	817	38.61 (5.345)	49.0	-2.75 (2.85)
			11935			

23andMe

Study Name	Origin	Study design	n	mean age of onset of myopia (sd)	% female
23andMe_V2	United States	population based study	12128	13.6 (5.8)	45.9
23andMe_V3	United States	population based study	92165	<10	45.3
			104293		

Supplementary Table 1b: Phenotyping and imputation methods per cohort

CREAM-EUR	Study Name	Phenotyping method	Genotyping chip	Pre-imputation QC metrics (exclusion criteria)	*N* excluded individuals Post-QC	Prephasing	Imputation method	GWAS software
				For MAF 0-0.05: SNP call rate of < 99%; for MAF >=0.05 SNP: call rate < 95%; p_HWE < 1x10-6; Sex	0	ShapeIt	IMPUTE2	SNPtest
1	1958 British Birth Cohort	Autorefractor Nikon Retinomax 2	Illumina Human1M-Duo Beadchip	disconcordance with reported, disconcordance with know controls/ repeats				
				SNPs: MAF <= 0.01, p_HWE <= 1e-6, SNP Call rate <= 0.98.				
	ALIENOR	Autorefractor Luneau SPEEDY K	Illumina Human 610-Quad BeadChip	Individuals: exclusion based on missingness + exclusion of duplicated/related individuals, individuals	20	SHAPEIT2	IMPUTE2	SNPtest
2				with discordance between clinical and genetic sex, or individuals with evidence of non-European				
3	ALSPAC-Mothers	Autorefractor Canon R50	Illumina 660W-Quad BeadChip	MAF <0.01, p_HWE <1x10-7, SNP call rate <95%	8196	SHAPEIT2	Minimac	mach2qtl
				MAF <0.01, p_HWE <5x10-10, SNP call rate <97%, individual call rate<97%. Identity by descent was				
				computed in PLINK based on autosomal markers, with one of each pair of individuals with relatedness				
				of > 0.2 removed. Participants with PC1 or PC2 values > 6 standard deviations from the known northern				
4	ANZRAG	Refractive details were obtained from clinical notes	Illumina Omni 1M/Illumina Omni Express	European ancestry group were excluded.	24	IMPUTE2	IMPUTE2	SNPtest
				MAF <0.02, p_HWE <1x10-4, SNP call rate <98%, individual call rate < 98%, used Illumina annotation to				
5	AREDS	Subjective Refraction	Illumina HumanOmni2.5-4v1_H array	identify and flip all the SNPs where the TOP allele was not on the "+" strand	0	SHAPEIT2	IMPUTE2	Plink
7	BMES	Autorefractor Zeiss Humphrey-530	Illumina 670 Quad Custom Chip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 95%	227	MACH	Minimac	ProbABEL
6	BATS	Autorefractor Zeiss Humphrey-598	Illumina HumanHap 610-Quad array	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	0	MACH	Minimac	MERLIN
8	Croatia-Korcula	Auto Ref/Keratometry NIDEK ARK30	Illumina 370CNV-Quad v1 BeadChip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	92	SHAPEIT2	IMPUTE2	ProbABEL
9	Croatia-Split	Auto Ref/Keratometry NIDEK ARK30	Illumina 370CNV-Quad v3 BeadChip and OMNI	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	41	SHAPEIT2	IMPUTE2	ProbABEL
10	Croatia-Vis	Auto Ref/Keratometry NIDEK ARK30	Illumina HumanHap300 v1.0 BeadChip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	65	SHAPEIT2	IMPUTE2	ProbABEL
11	DCCT	Subjective Refraction	Illumina Human-1M BeadChip	Sample QC: gender mismatch with typed X-linked markers (n=3), call rate < 0.95 (n=0), genotype	113	SHAPEIT2	IMPUTE2	SNPtest
				discrepancy with an earlier study (n=58), autosomal heterozygosity > 0.32 (n=0), cryptic relatedness				
				(n=2), self-reported ethnicity other than white (n=50), outliers in PCA (n=24)				
12	EGCUT	Eye glass prescriptions	Illumina Omni Express	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	0	IMPUTE2	IMPUTE2	SNPtest
				SNP QC: callrate (95% minimum, male-specific on Y), cluster pattern (using Affymetrix SNPolisher),				
				plate effect on minor allele frequency, Hardy Weinberg p < 1e-8. Sample QC: CEL file not generated				
				(raw image quality to poor to process), DishQC < 0.82 (poor fluorescence signal contrast), Step 1 call				
				rate < 97% (calculated on a subset of SNPs, intended to ensure clustering uses only good quality data),				
13	EPIC-Norfolk	Autorefractor Zeiss Humphrey-500	Affymetrix GeneChip Human Mapping 500K Array Set	Sample identity unclear (lab team flagged, or unexpected duplicate or non-duplicate), Sex	N/A	SHAPEIT2	IMPUTE2	SNPtest
				discordance or other sex chromosome abnormality detected, Final sample callrate < 97%,				
				Heterozygosity outlier (calculated separately for SNPs with minor allele frequency >1% and <1%), Rare				
				allele count outlier (unusually high number of singleton, doubleton or tripleton counts), For				
				duplicate/triplicate samples, only the sample with the highest call rate was kept, After this cleaning,				
14	ERF	Topcon RM-A2000 autorefractor	Illumina 6K, 318K, 350K, 610K, Affymetrix 250K	MAF <0.005, p_HWE < 1x10-8, SNP call rate < 98%	1048	MACH	Minimac	ProbABEL
15	FECD	Subjective Refraction	HumanOmni2.5-Quad BeadChip	MAF < 0.02, p_HWE < 1x10-4 SNP call rate < 98%, individual call rate < 98%, > 1 Mendelian error	0	SHAPEIT2	IMPUTE2	Plink
16	FITSA	TOPCON AT	Illumina HumanCoreExome	MAF < 0.01, p_HWE < 1x10-6, SNP call rate < 95%	4	SHAPEIT2	IMPUTE2	SNPtest
17	Framingham	Subjective Refraction	Affymetrix Mapping500K (Nsp & Sty) + 50K HumanHap	Individual-level QC: Non-caucasian by history, individual missing genotypes > 0.10, mendelian	N/A	SHAPEIT2	IMPUTE2	MixABEL
			supplement	inconsistencies > 0.1,; Marker-level QC: MAF < 0.01, SNP call rate < 97%, p_HWE in unrelateds < 10e-05;				
				Note: Baseline QC was carried out on all genotyped Framingham Study Participants (N=9,270) in order				
				maximize the accuracy of marker-level and family-level statistics.				
18	Gutenberg Health Study 1	Zeiss Humphrey® Automated Refractor/Keratometer (HARK) 599™	Affymetrix Genome-Wide Human SNP Array 6.0	MAF ≤0.01, p_HWE≤0.0001, SNP call rate ≤ 98%	N/A	IMPUTE2	IMPUTE2	SNPtest
19	Gutenberg Health Study 2	Zeiss Humphrey® Automated Refractor/Keratometer (HARK) 599™	Affymetrix Genome-Wide Human SNP Array 6.0	MAF ≤0.01, p_HWE≤0.0001, SNP call rate ≤ 98%	N/A	IMPUTE2	IMPUTE2	SNPtest
20	KORA	Nikon Retinomax and Eye glass prescriptions	Illumina Omni 2.5/Illumina Omni Express	MAF < 0.01, p_HWE < 1x10-6, SNP call rate < 98%	N/A	SHAPEIT2	IMPUTE2	SNPtest
21	OGP Talana	Autorefractor Topcon RK-8100	Affymetrix GeneChip Human Mapping 500K	MAF <0.05, p_HWE < 1x10-4, SNP call rate < 95%	0	SHAPEIT2	IMPUTE2	ProbABEL
22	ORCADES	KOWA	Illumina HumanHap 300v2, 370CNV-Quad and OMNI	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	98	SHAPEIT2	IMPUTE2	ProbABEL
23	Rotterdam Study I	Topcon RM-A2000 autorefractor	Illumina HumanHap 550-Duo v3 BeadChip, Illumina HumanHap 610-Quad BeadChip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	504	MACH	Minimac	ProbABEL
24	Rotterdam Study II	Topcon RM-A2000 autorefractor	Illumina HumanHap 550-Duo v3 BeadChip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	119	MACH	Minimac	ProbABEL
25	Rotterdam Study III	Topcon RM-A2000 autorefractor	Illumina HumanHap 610-Quad BeadChip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	98	MACH	Minimac	ProbABEL
26	TEST	Autorefractor Zeiss Humphrey-598	Illumina HumanHap 610-Quad BeadChip	MAF <0.01, p_HWE < 1x10-6, SNP call rate < 98%	0	MACH	Minimac	MERLIN
27	Twins UK	ARM-10 Autorefractor (Takagi Ltd)	Illumina 610K, Illumina 317K	For MAF 0-0.05: SNP call rate of < 99%; for MAF >=0.05 SNP: call rate < 95%; p_HWE < 1x10-6;			IMPUTE2	SNPtest
				Heterozygosity within 3SD (individuals); Sex disconcordance with reported, disconcordance with know				
28	WESDR	Subjective Refraction	Illumina Omni1-Quad BeadChip	Sample QC: gender mismatch with typed X-linked markers (n=9), cryptic relatedness (n=24),	58	SHAPEIT2	IMPUTE2	SNPtest
				autosomal heterozygosity > 0.3 (n=5), call rate < 0.95 (n=29), self-reported ethnicity other than white				
				(n=3), outliers in MDS analysis (n=4); SNP QC: duplicate SNPs, high missing rate [maf>=0.05 &				
				MISS>0.05, 0.05>MAF>=0.01 & MISS>0.01, MAF<0.01 & MISS>0], p_HWE<1E-9				
29	YFS	NIDEK AR-310AR autorefractor	Illumina 670K Custom Array	MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%	963	SHAPEIT1	IMPUTE2	SNPtest

CREAM-ASN	Study Name	phenotyping method	GWAS chip	QC parameters genotypes	Post-QC	Prephasing	Imputation method	GWAS software
1	Beijing Eye Study	Canon RK-5 Auto Ref-Keratometer	Illumina Human 610 Quad Beadchip	MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%, Samples with cryptic relatedness, populati	156	MACH	Minimac	SNPtest
2	Nagahama	Nidek-ARK530A	Illumina HumanHap 610-Quad BeadChip, Illumina Human Omni2.5-8 BeadChip, Illumina Human Omni2.5-Quad BeadChip, Illumina Human Omni2.5S-8 BeadChip, Illumina Infinium Exome-24 v1.0 BeadChip			SHAPEIT2	Minimac	Plink
				MAF<0.01, p_HWE<1x10-7, SNP call rate<90%, individual call rate<90%	0			
				MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%, Samples with cryptic relatedness,				
3	SCES	Canon RK-5 Auto Ref-Keratometer	Illumina Human 610 Quad Beadchip	population structure ascertainment and excessive heterogeneity	63	MACH	Minimac	SNPtest
4	SIMES	Canon RK-5 Auto Ref-Keratometer	Illumina Human 610 Quad Beadchip	MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%	530	MACH	Minimac	SNPtest
				MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%,Samples with cryptic relatedness,				
5	SINDI	Canon RK-5 Auto Ref-Keratometer	Illumina Human 610 Quad Beadchip	population structure ascertainment and excessive heterogeneity	415	MACH	Minimac	SNPtest
				MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%,Samples with cryptic relatedness,				
6	SP2-1M	Canon RK-5 Auto Ref-Keratometer	Illumina Human1M-Duo v3 BeadChip	population structure ascertainment and excessive heterogeneity	63	MACH	Minimac	SNPtest
				MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%,Samples with cryptic relatedness,				
7	SP2-610	Canon RK-5 Auto Ref-Keratometer	Illumina Human 610 Quad Beadchip	population structure ascertainment and excessive heterogeneity	321	MACH	Minimac	SNPtest
				MAF <0.01, p_HWE < 1x10-6, Sample and SNP call rate < 95%, samples with >1% SNPs showing				
8	STARS	Canon RK-F1 Autorefractor, Welch Allyn retinoscopy	Illumina Human 610 Quad Beadchip	Mendelian error,Samples with cryptic relatedness, population structure ascertainment and excessive	26	MACH	Minimac	SNPtest

23andMe	Study Name	phenotyping method	GWAS chip	QC parameters genotypes	Post-QC	Prephasing	Imputation method	GWAS software
1	23andMe_V2	Questionaire	Illumina!HumanHap550+ BeadChip	MAF < 0.001, individuals who have <97% European ancestry, p_HWE < 1x10-20, SNP call rate < 95%, or	N/A	Beagle	Minimac	Cox proportional hazards model using R and custom GWAS software
				with large allele frequency discrepancies compared to the 1000 Genomes reference data				
2	23andMe_V3	Questionaire	OmniExpress+ BeadChip	MAF < 0.001, individuals who have <97% European ancestry, p_HWE < 1x10-20, SNP call rate < 95%, or	N/A	Beagle	Minimac	Cox proportional hazards model using R and custom GWAS software
				with large allele frequency discrepancies compared to the 1000 Genomes reference data				

Supplementary Table 5: Index SNPs HapMap II from CREAM and 23andMe

SNP information				HAPMAPII CREAM and 23andMe topSNPs for Refractive Error and Age of Diagnosis Myopia								1000G meta-analysis results of CREAM and 23andMe topSNPs from HapMapII for Refractive Error and Age of Diagnosis Myopia									
Locus #	rsnumber	CHR	POSITION	A1/A2	Locus Name CREAM	CREAM Beta	CREAM SEM	CREAM P value	Locus Name 23andMe	23andMe HR (CI)	23andMe P value	A1	Freq1	Zscore	P-value	Direction	HetISq	HetChi	HetDf	HetPval	N
1	rs1652333	1	207470460	G/A	*CD55*	-0.1116533	0.0160072	3.05434E-12	-	-	-	G	0.6769	-7.242	4.42E-13	++	81.7	5.457	1	0.01949	160136
2	rs1656404	2	233379941	A/G	*PRSS56*	-0.1528567	0.0235048	7.86186E-11	-	-	-	A	0.1733	-11.253	2.25E-29	--	75	3.999	1	0.04554	156056
	rs1550094		233385396	A/G	-	-	-	-	*PRSS56*	1.087 (1.067 - 1.107)	5.8E-18	A	0.7005	12.738	3.64E-37	++	88.9	8.996	1	0.002705	159422
3	rs17400325	2	178565913	T/C	-	-	-	-	*PDE11A*	1.144 (1.099 - 1.190)	8.7E-11	T	0.9513	7.994	1.30E-15	++	92.1	12.616	1	0.0003824	150322
4	rs17412774	2	146773948	A/C	-	-	-	-	*PABPCP2*	0.933 (0.917 - 0.950)	1.5E-14	A	0.5506	-8.568	1.05E-17	--	88.9	9.033	1	0.002652	159506
5	rs17428076	2	172851936	C/G	-	-	-	-	*DLX1*	0.935 (0.916 - 0.955)	1.4E-10	C	0.768	-8.183	2.77E-16	--	88.8	8.899	1	0.002854	160151
6	rs1881492	2	233406998	T/G	*CHRNG*	-0.139	0.021	5.15E-11	-	-	-	T	0.2008	-10.252	1.16E-24	--	0	0.338	1	0.5611	159506
7	rs13091182	3	141133960	G/A	-	-	-	-	*ZBTB38*	0.940 (0.923 - 0.958)	3.6E-11	G	0.3352	-8.001	1.23E-15	--	83.8	6.166	1	0.01303	153193
8	rs14165	3	53847408	A/G	*CACNA1D*	0.096	0.017	2.14E-08	-	-	-	A	0.3014	6.25	4.10E-10	++	0	0.061	1	0.8045	149655
9	rs5022942	4	81959966	G/A	-	-	-	-	*BMP3*	1.076 (1.054 - 1.098)	4.2E-12	G	0.2416	9.258	2.08E-20	++	89.8	9.848	1	0.0017	160150
10	rs9307551	4	80530671	A/C	*LOC100506035*	-0.099	0.017	1.09E-08	-	-	-	A	0.2452	-7.972	1.56E-15	--	0	0.237	1	0.6265	160149
11	rs12205363	6	129834629	C/T	*LAMA2*	0.235	0.033	1.79E-12	-	-	-	C	0.9317	15.975	1.91E-57	++	96.6	29.438	1	5.78E-08	150327
	rs12193446		129820038	A/G	-	-	-	-	*LAMA2*	0.788 (0.763 - 0.813)	6.8E-53	A	0.9063	-19.431	4.21E-84	--	98.4	60.996	1	5.72E-15	150269
12	rs7744813	6	73643289	A/C	*KCNQ5*	-0.112	0.019	4.18E-09	-	-	-	A	0.5905	-14.555	5.43E-48	--	90.6	10.598	1	0.001132	160091
				A/C	-	-	-	-	*KCNQ5*	0.909 (0.893 - 0.926)	2.7E-25										
13	rs7837791	8	60179086	T/G	*TOX*	0.106	0.015	3.99E-12	-	-	-	T	0.4816	11.59	4.64E-31	++	53	2.127	1	0.1447	160152
	chr8:60178580		60178580	C/G	-	-	-	-	*TOX/CA8*	0.914 (0.897 - 0.931)	4E-22	C	0.6415	-13.137	2.03E-39	--	87	7.678	1	0.00559	160128
14	rs4237036	8	61701057	C/T	*CHD7*	0.089	0.016	1.82E-08	-	-	-	C	0.6669	5.205	1.94E-07	++	0	0.984	1	0.3212	160148
15	rs7829127	8	40726394	A/C	*ZMAT4*	-0.116	0.018	3.69E-10	-	-	-	A	0.7917	-10.911	1.02E-27	--	92.4	13.217	1	0.0002774	160132
	rs7829127			A/G	-	-	-	-	*SFRP1*	0.901 (0.880 - 0.923)	1.8E-18										
16	rs11145465	9	71766593	A/C	*TJP2*	-0.124	0.021	7.26E-09	-	-	-	A	0.2122	-9.546	1.35E-21	--	46.3	1.863	1	0.1722	153174
	rs11145746		71834380	G/A	-	-	-	-	*TJP2*	1.087 (1.063 - 1.112)	5.2E-13	G	0.2056	9.098	9.22E-20	++	68	3.127	1	0.07698	153113
17	rs7042950	9	77149837	G/A	*RORB*	-0.0964935	0.0175941	4.14842E-08	-	-		G	0.7323	-6.797	1.07E-11	--	0	0.012	1	0.9122	160153
18	rs10882165	10	94924324	T/A	*CYP26A1*	-0.107	0.016	1.03E-11	-	-	-	T	0.5869	-6.155	7.49E-10	--	69.1	3.237	1	0.07198	155329
19	rs6480859	10	79081948	C/T	-	-	-	-	*KCNMA1*	1.058 (1.039 - 1.077)	7.3E-10	C	0.363	8.202	2.36E-16	++	79	4.765	1	0.02904	160148
20	rs7084402	10	60265404	G/A	*BICC1*	-0.108	0.015	2.06E-13	-	-	-	G	0.5277	-8.828	1.07E-18	--	0	0.428	1	0.5129	160020
21	rs745480	10	85986554	C/G	-	-	-	-	*RGR*	1.063 (1.044 - 1.081)	8E-12	C	0.5109	8.314	9.26E-17	++	67.3	3.055	1	0.0805	159504
22	rs11601239	11	105556598	C/G	*GRIA4*	-0.0949272	0.0163137	5.92475E-09	-	-	-	C	0.485	-6.824	8.84E-12	--	0	0.008	1	0.9281	160118
23	rs1381566	11	40149607	T/G	-	-	-	-	*LRRC4C*	1.149 (1.122 - 1.176)	2.3E-30	T	0.81	13.593	4.43E-42	++	97.6	40.832	1	1.66E-10	157519
24	rs2155413	11	84634790	C/A	-	-	-	-	*DLG2*	1.061 (1.043 - 1.080)	1.7E-11	C	0.4823	7.755	8.85E-15	++	92.4	13.078	1	0.0002987	159504
25	rs12229663	12	71249996	G/A	*PTPRR*	0.099	0.017	5.47E-09	-	-	-	G	0.7527	7.362	1.81E-13	++	0	0.221	1	0.6386	160133
26	rs3138142	12	56115585	C/G	*RDH5*	0.119	0.017	4.44E-12	-	-	-	T	0.2115	13.766	4.06E-43	++	95.9	24.566	1	7.18E-07	157544
				C/G	-	-	-	-	*RDH5*	0.890 (0.870 - 0.911)	1.3E-23										
27	rs2184971	13	100818092	G/A	*PCCA*	0.085	0.015	2.11E-08	-	-	-	G	0.5441	5.232	1.68E-07	++	0	0.51	1	0.4751	160146
28	rs4291789	13	100672921	C/G	-	-	-	-	*ZIC2*	1.069 (1.046 - 1.092)	2.10E-08	C	0.6723	7.899	2.80E-15	++	86	7.146	1	0.007514	159988
29	chr14:54413001	14	54413001	G/C	-	-	-	-	*BMP4*	0.946 (0.929 - 0.963)	1.1E-09	G	0.4657	-7.118	1.09E-12	--	73.6	3.784	1	0.05174	160104
30	rs1254319	14	60903757	A/G	*SIX6*	-0.088	0.015	1.00E-08	-	-	-	A	0.3322	-5.602	2.12E-08	--	76.3	4.226	1	0.03981	160153
31	rs4778879	15	79372875	G/A	*RASGRF1*	-0.102	0.015	4.25E-11	-	-	-	G	0.5823	-9.898	4.24E-23	--	0	0.684	1	0.4081	160068
	rs28412916		79378167	A/C	-	-	-	-	*RASGRF1*	1.067 (1.048 - 1.086)	8.2E-13	A	0.5874	9.944	2.68E-23	++	0	0.161	1	0.6886	160152
32	rs524952	15	35005886	T/A	*GJD2*	0.1582561	0.019821	1.44329E-15	-	-	-	T	0.4748	17.075	2.28E-65	++	83	5.865	1	0.01544	160150
				T/A	-	-	-	-	*GOLGA8B/GJD2*	1.089 (1.070 - 1.108)	6.9E-22										
33	rs17648524	16	7459683	G/C	*A2BP1*	0.118	0.019	5.64E-10	-	-	-	G	0.3535	13.693	1.13E-42	++	96.3	26.829	1	2.22E-07	160122
			7459683	G/C	-	-	-	-	*RBFOX1*	1.102 (1.082 - 1.122)	4.1E-26										
34	rs17183295	17	31078272	T/C	*MYO1D*	-0.131	0.02	9.66E-11	-	-	-	T	0.1901	-8.177	2.91E-16	--	0	0.41	1	0.522	152597
35	rs2908972	17	11407259	T/A	*SHISA6*	0.101	0.015	7.29E-11	-	-	-	T	0.4146	11.125	9.46E-29	++	23	1.299	1	0.2544	160123
				T/A	-	-	-	-	*SHISA6*	1.074 (1.055 - 1.093)	5.2E-15										
36	rs4793501	17	68718734	C/T	*KCNJ2*	0.08040793	0.0144771	2.78971E-08	-	-	-	C	0.5748	7.212	5.53E-13	++	0	0.288	1	0.5917	160150
37	rs235770	20	6761765	T/C	*BMP2*	-0.089	0.016	1.57E-08	-	-	-	T	0.3717	-5.926	3.11E-09	--	0	0.362	1	0.5474	157521

Supplementary Table 6. LD score regression analysis and heritability explained by common SNPs

Sample	Ref. panel[1]	N[2]	λ_{GC}	LD score regression[3]		
				Intercept (se)	h^2 (se)	p
CREAM-EUR	EUR	44.192	1.165	1.023 (0.008)	0.214 (0.015)	2.4 x 10e-49
23andMe	EUR	104.292	1.108	0.892 (0.009)	0.172 (0.009)	1.6 x 10e-81
CREAM-EAS	EAS	9.826	1.017	1.001 (0.006)	0.053 (0.040)	0.190

[1]Reference panel used to calculate LD scores.

[2]The maximum sample size; not all markers were genotyped in all samples and therefore this number will vary from marker to marker.

[3]LD score regression intercept and heritability estimate (p-value relates to a test of the null hypothesis of $h^2=0$; intercepts above 1 give evidence of inflated association statistics due to residual population stratification or relatedness.

Supplementary Table 7: Exonic protein-altering variants

Marker name	Gene name	P-value meta-analysis Stage 3	chr:pos	Mutation	Amino acid change
rs5442	*GNB3*	5.48E-15	12:6954864	missense	Gly > Ser
rs807037	*KAZALD1*	4.59E-08	10:102824349	missense	Gly > Ala
rs1550094	*PRSS56*	3.64E-37	2:233385396	missense	Ala > Thr
rs1064583	*COL10A1*	6.90E-11	6:116446576	missense	Met > Thr
rs6420484	*TSPAN10*	2.80E-09	17:79612397	missense	Tyr > Cys
rs2303635	*AMOTL2*	1.14E-09	3:134086356	missense	Ala > Pro
rs35337422	*RD3L*	2.19E-10	14:104407243	missense	Ile > Arg
rs17400325	*PDE11A*	1.30E-15	2:178565913	missense	Tyr > Cys

P values derived from stage 3 meta-analysis (n=160,420)

***** This locus did not replicate in UKEV

Supplementary Table 8: Post-GWAS analyses additional genes

Gene	Chr	Start	End	N SNPs per region tested	P value stage 3 meta-analysis	Analysis	Remarks	Novel gene #
MIR6512	2	178178533	178178610	72	4.52E-07	fastBAT		1
RPP14	3	57062260	58625174	5003	4.46E-07	fgwas		2
IL4	5	132009678	132018368	1	1.70E-06	Eugene		3
HMGN4	6	26538633	26546482	7	1.00E-07	Eugene, fgwas		4
HIST1H2AG	6	27100816	27101314	58	1.38E-06	fastBAT		5
CDKAL1	6	19752146	21152237	5010	6.97E-08	fgwas	Genome wide significant in Discovery stage	
POU6F2	7	39017608	39504390	526	5.33E-07	fastBAT		6
CHD7	8	61591323	61780586	248	5.36E-07	fastBAT	Significant in Discovery stage after Bonferroni's correction and established in HapMap II analysis in 2013	7
C8orf84	8	73887862	75301149	5019	2.87E-07	fgwas	same locus 1	8
BC127738	8	73887862	75301149	5019	2.87E-07	fgwas	same locus 1	9
TRAF1	9	123664671	123691451	3	1.20E-06	Eugene		10
C5	9	123714613	123812554	109	1.10E-06	fastBAT		11
C10orf11	10	77542518	78317126	670	1.14E-06	fastBAT	Genome wide significant in Discovery stage	
TLX1	10	102891060	102897546	132	5.88E-07	fastBAT, fgwas		12
ACP2	11	47260853	47270457	4	1.10E-06	Eugene		13
SYTL2	11	85405267	85522184	4	1.60E-06	Eugene		14
CCDC89	11	85394892	85397320	74	5.50E-07	fastBAT		15
HNRNPKP3	11	43283053	43290919	115	4.72E-07	fastBAT	Genome wide significant in Discovery stage	
TP53AIP1	11	128804626	128813294	174	1.73E-07	fastBAT		16
ANKRD9	14	102973179	102976136	6	4.00E-07	Eugene		17
CRHR1-IT1	17	4369769	43725582	3	1.00E-07	Eugene		18
LOC101927557	17	55599608	55600958	190	1.12E-06	fastBAT		19
TNFSF12	17	7452374	7461207	117	1.82E-06	fastBAT	Genome wide significant in Discovery stage	
FAM83C	20	33873533	33880225	71	1.67E-06	fastBAT		20
TMEM184B	22	38615297	38669040	164	1.72E-06	fastBAT		21
LARGE	22	32001050	33156201	5019	4.13E-07	fgwas	same locus 2	22
ISX	22	32001050	33156201	5019	4.13E-07	fgwas	same locus 2	23
LOC388906	22	39756985	41941243	5019	2.00E-07	fgwas	same locus 3	24
BC038245	22	39756985	41941243	5019	2.00E-07	fgwas	same locus 3	25

P values derived from stage 3 meta-analysis (n=160,420)

Supplementary Table 9: Predictive power of the polygenic score in Rotterdam Studies (I, II, III)

Score	MEGA GWAS P-value threshold*	*n* variants in score	*R2*	P-value	% variance explained†
Reference**			0.052	NA	5.2
S1	5.00E-08	152	0.099	5.35E-122	4.7
S2	5.00E-07	214	0.101	1.17E-127	4.9
S3	5.00E-06	334	0.106	2.24E-139	5.4
S4	5.00E-05	661	0.112	2.33E-155	6.0
S5	5.00E-04	1815	0.122	1.91E-181	7.0
S6	0.005	7303	0.130	1.19E-203	7.8
S7	0.01	11763	0.129	2.20E-201	7.7
S8	0.05	37999	0.124	2.91E-188	7.2
S9	0.1	63106	0.119	8.93E-176	6.8
S10	0.5	183871	0.113	1.32E-159	6.2
S11	0.8	228198	0.113	3.25E-158	6.1
S12	1	243938	0.113	5.53E-158	6.1

* Meta-analysis results stage 3 (CREAM and 23andME) excluding RS-I-II-III (n=148,645)

**Reference model, mean Spherical Equivalent ~ age + sex + first 5PCs + cohort

† Variance explained by each score (S1-S12) was calculated by substraction of a full model (reference + score) from a reduced model (reference only)

Estimating the incremental *R2* from including the polygenic score (S1-S12) in a regression of mean spherical equivalent on age, sex, first 5 principal components and cohort

Supplementary Table 12: RNA genes – region look-ups

	Alias	Class of RNA	Chromosome:Position	Genetic variant (Stage 3 meta-	Description
5S_rRNA	5S ribosomal RNA	Ribosomal RNA	1:163479274-163479385	rs1556867	
AK097193/LOC101926964	AK097193: cDNA - transcript of LOC101926964: alias RP11-436K8.1	Uncharacterized non-protein coding RNA	1:61125303-61291256	rs11589487	
AK097934/LINC01237**	cDNA clone - transcript of LINC01237	Long intergenic non-protein coding RNA	2:242912834-242919427	rs12998513	LOC102723927 also in region
AK123891/FLJ41897	cDNA clone - transcript of FLJ41897	Uncharacterized non-protein coding RNA	22:32896532-32898414	rs9606967	
AK124857/LOC101927394	cDNA clone - transcript of LOC101927394	Uncharacterized non-protein coding RNA	3:7994492-8057994	rs9681162	
BC030753/NFIA-AS2	cDNA clone - transcript of NFIA-AS2 (alias: NFIA Antisense RNA 2)	Uncharacterized non-protein coding RNA	1:61405916-61436448	rs11589487	
BC035400/LOC102723409	cDNA clone - transcript of LOC102723409	Uncharacterized non-protein coding RNA	6:129800758-129873731	rs12193446	Overlap with LAMA2 region
BC039327/CASC17	cDNA clone - transcript of CASC17 (alias: cancer susceptibility 17)	Long intergenic non-protein coding RNA	17:69093915-69198318	rs4793501	Diseases associated with CASC17 include Prostate Cancer Susceptibility (genecards).
BC040861	cDNA clone	Uncharacterized non-protein coding RNA	2:145780382-145910072	rs56075542	
BC127738/LOC100130301	cDNA clone - transcript of LOC100130301	Uncharacterized non-protein coding RNA	8:74153659-74171737	Gene based (rs16938625 closest GWS tophit)	
CRHR1-IT1	Long Intergenic Non-Protein Coding RNA 2210; C17orf69	Long intergenic non-protein coding RNA	17:43716341-43723595	Gene based (rs117118311 closest GWS tophit)	
D43770/LINC01152	RNA of LINC01152	Long intergenic non-protein coding RNA	17:70026957-70035822	rs7207217	This RNA may play a significant role in differentiation or sex determination (NCBI)
FLJ16171/LINC01951	LINC01951	Long intergenic non-protein coding RNA	5:174346085-174422734	rs7449443	
HP08777	RP3-522D1.1	Uncharacterized non-protein coding RNA	1:113392522-113420491	rs1237670	Overlap with LINC01356 region
LINC00333	Long Intergenic Non-Protein Coding RNA 00333	Long intergenic non-protein coding RNA	13:84714737-85180903	rs9547035	
LINC00340/CASC15	CASC15, Cancer Susceptibility 15 (Non-Protein Coding); Lnc-SOX4-1; Long intergenic non-protein coding RNA 340	Long intergenic non-protein coding RNA	6:22146883-22194616	rs1207782	
LINC00351	Long Intergenic Non-Protein Coding RNA 00351	Long intergenic non-protein coding RNA	3:85937738-86118797	rs9547035	
LINC00461	long intergenic non-protein coding RNA 461; EyeLinc1; Visual Cortex Expressed; ECONEXIN	Long intergenic non-protein coding RNA	5:87836597-87980620	rs7737179	Highly conserved and expressed in the human macula (PMID: 23562822). ECONEXIN was the most highly conserved intergenic lncRNA containing 83.0% homology with the mouse ortholog (C130071C03Rik) for a region over 2500 bp in length within its exon 3. Expressions of ECONEXIN and C130071C03Rik were significantly upregulated in both human and mouse glioma tissues. PMID: 28368417
LINC00862	C1orf98; SMIM16;Small Integral Membrane Protein 16; Long Intergenic Non-Protein Coding RNA 862	Long intergenic non-protein coding RNA	1:200311672-200342920	rs2225986	
LMCD1-AS1	LMCD1 antisense RNA 1	Uncharacterized non-protein coding RNA	3:8235936-8543344	rs9681162	
LOC100506035/LINC00989	Long Intergenic Non-Protein Coding RNA 00989	Uncharacterized non-protein coding RNA	4:80413747-80497614	rs7662551	
LOC100508120/GMDS-AS1	GMDS antisense RNA 1	Uncharacterized non-protein coding RNA	6:2245987-2413825	rs10458138	
LOC101927557	Long Intergenic Non-Protein Coding RNA 101927557	Uncharacterized non-protein coding RNA	17:55599609-55600958	Gene based (rs28488643 closest GWS tophit)	
MIR6512	MicroRNA 6512	MicroRNA	2:178178534-178178610	Gene based (rs2573081 closest GWS tophit)	
SNORA40	Small nucleolar RNA 40, H/ACA box 40	Small nucleolar RNA	2:16380471-16380563	rs28658452	
SNORA51	Small nucleolar RNA 51, H/ACA Box 51; ACA51 SnoRNA; ACA51	Small nucleolar RNA	8:60049931-60050061	rs72621438	
TMEM161B-AS1	TMEM161B Antisense RNA 1 (Non-Protein Coding)	Uncharacterized non-protein coding RNA	5:87564699-87732491	rs7737179	
TRNA_Ala	Alanine tRNA	transfer RNA	14:89445442-89445514	rs17125093	
TRNA_Gln	Glutamine tRNA	transfer RNA	17:47269890-47269961	rs11654644	
TRNA_Ser	Serine tRNA	transfer RNA	6:28180815-28180896	rs1150687	
U6	RNA, U6 Small Nuclear	Small nucleolar RNA	4:83095700-83095806	rs2166181	Diseases associated with RNU6-1 include Poikiloderma With Neutropenia. Among its related pathways are mRNA Splicing - Major Pathway and RNA transport. (genecards)

** These loci were not included in the UKEV replication analysis

Supplementary Table 17: Genes in Dopamine pathway – region look-ups

Gene	Chromosome	Start	End	NCBI Reference Sequence
DRD1	5	174867675	174871163	NC_000005.9
DRD2	11	113280317	113346413	NC_000011.9
DRD3	3	113847499	113918254	NC_000003.11
DRD4	11	637305	640706	NC_000011.9
DRD5	4	9783258	9785633	NC_000004.11
COMT	22	19929263	19957498	NC_000022.10
DBH	9	136501485	136524466	NC_000009.11
DDC	7	50526134	50633154	NC_000007.13
MAOA	X	43514155	43606071	NC_000023.10
TH	11	2185159	2193107	NC_000011.9
SLC6A3/DAT	5	1392905	1445543	NC_000005.9
SLC6A4/SERT	17	28521337	28562986	NC_000017.10

Gene	Chr	Start	End	N_SNPs tested	Lowest P-value Stage 3 meta-analysis	TopVariant	Position Top Variant
COMT	22	19929263	19957498	3971	2.00E-04	rs4819854	19854659
DBH	9	136501485	136524466	5519	2.78E-04	rs191849948	136719162
DDC	7	50526134	50633154	4230	2.38E-04	rs2189432	51100959
DRD1	5	174867675	174871163	3733	3.58E-08	rs7449443	174720893
DRD2	11	113280317	113346413	3988	1.07E-04	rs188263557	113108780
DRD3	3	113847499	113918254	2915	9.41E-04	rs189807093	114022413
DRD4	11	637305	640706	4760	1.01E-04	rs80190876	680614
DRD5	4	9783258	9785633	5705	1.03E-04	rs57553236	9688140
MAOA	X	43514155	43606071	1591	5.76E-03	rs73196113	43972237
SLC6A3/DAT	5	1392905	1445543	5846	2.35E-04	rs147426622	1652855
SLC6A4/SERT	17	28521337	28562986	2426	4.77E-04	rs113822901	28994581
TH	11	2185159	2193107	4641	3.24E-04	rs112121578	1937116

References

1. Rahi, J.S., Cumberland, P.M. & Peckham, C.S. Myopia over the lifecourse: prevalence and early life influences in the 1958 British birth cohort. *Ophthalmology* **118**, 797-804 (2011).
2. Delcourt, C. *et al.* Nutrition and age-related eye diseases: the Alienor (Antioxydants, Lipides Essentiels, Nutrition et maladies OculaiRes) Study. *J Nutr Health Aging* **14**, 854-61 (2010).
3. Group, C.S. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* **22**, 316-25 (2003).
4. Fraser, A. *et al.* Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* **42**, 97-110 (2013).
5. Boyd, A. *et al.* Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* **42**, 111-27 (2013).
6. Age-Related Eye Disease Study Research, G. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E and beta carotene for age-related cataract and vision loss: AREDS report no. 9. *Arch Ophthalmol* **119**, 1439-52 (2001).
7. Age-Related Eye Disease Study Research, G. A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol* **119**, 1417-36 (2001).
8. Clemons, T.E., Chew, E.Y., Bressler, S.B., McBee, W. & Age-Related Eye Disease Study Research, G. National Eye Institute Visual Function Questionnaire in the Age-Related Eye Disease Study (AREDS): AREDS Report No. 10. *Arch Ophthalmol* **121**, 211-7 (2003).
9. Age-Related Eye Disease Study Research, G. The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. *Control Clin Trials* **20**, 573-600 (1999).
10. Louttit, M.D. *et al.* A multicenter study to map genes for Fuchs endothelial corneal dystrophy: baseline characteristics and heritability. *Cornea* **31**, 26-35 (2012).
11. Mitchell, P., Smith, W., Attebo, K. & Wang, J.J. Prevalence of age-related maculopathy in Australia. The Blue Mountains Eye Study. *Ophthalmology* **102**, 1450-60 (1995).
12. Foran, S., Wang, J.J. & Mitchell, P. Causes of visual impairment in two older population cross-sections: the Blue Mountains Eye Study. *Ophthalmic Epidemiol* **10**, 215-25 (2003).
13. Vitart, V. *et al.* Heritabilities of ocular biometrical traits in two croatian isolates with extended pedigrees. *Invest Ophthalmol Vis Sci* **51**, 737-43 (2010).
14. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Engl J Med* **329**, 977-86 (1993).
15. Group, D.E.R. *et al.* Intensive diabetes therapy and ocular surgery in type 1 diabetes. *N Engl J Med* **372**, 1722-33 (2015).
16. Nelis, M. *et al.* Genetic structure of Europeans: a view from the North-East. *PLoS One* **4**, e5472 (2009).
17. Milani, L., Leitsalu, L. & Metspalu, A. An epidemiological perspective of personalized medicine: the Estonian experience. *J Intern Med* **277**, 188-200 (2015).
18. Riboli, E. & Kaaks, R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* **26 Suppl 1**, S6-14 (1997).
19. Day, N. *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *Br J Cancer* **80 Suppl 1**, 95-103 (1999).
20. Hayat, S.A. *et al.* Cohort profile: A prospective cohort study of objective physical and cognitive capability and visual health in an ageing population of men and women in Norfolk (EPIC-Norfolk 3). *Int J Epidemiol* **43**, 1063-72 (2014).
21. Khawaja, A.P. *et al.* The EPIC-Norfolk Eye Study: rationale, methods and a cross-sectional analysis of visual impairment in a population-based cohort. *BMJ Open* **3**(2013).
22. Aulchenko, Y.S. *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-34 (2004).

23. Pardo, L.M., MacKay, I., Oostra, B., van Duijn, C.M. & Aulchenko, Y.S. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* **69**, 288-95 (2005).
24. Parssinen, O. *et al.* Heritability of spherical equivalent: a population-based twin study among 63- to 76-year-old female twins. *Ophthalmology* **117**, 1908-11 (2010).
25. Leibowitz, H.M. *et al.* The Framingham Eye Study monograph: An ophthalmological and epidemiological study of cataract, glaucoma, diabetic retinopathy, macular degeneration, and visual acuity in a general population of 2631 adults, 1973-1975. *Surv Ophthalmol* **24**, 335-610 (1980).
26. Familial aggregation and prevalence of myopia in the Framingham Offspring Eye Study. The Framingham Offspring Eye Study Group. *Arch Ophthalmol* **114**, 326-32 (1996).
27. Wichmann, H.E., Gieger, C., Illig, T. & Group, M.K.S. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67 Suppl 1**, S26-30 (2005).
28. Oexle, K. *et al.* Novel association to the proprotein convertase PCSK7 gene locus revealed by analysing soluble transferrin receptor (sTfR) levels. *Hum Mol Genet* **20**, 1042-7 (2011).
29. Holle, R., Happich, M., Lowel, H., Wichmann, H.E. & Group, M.K.S. KORA--a research platform for population based health research. *Gesundheitswesen* **67 Suppl 1**, S19-25 (2005).
30. Biino, G. *et al.* Ocular refraction: heritability and genome-wide search for eye morphometry traits in an isolated Sardinian population. *Hum Genet* **116**, 152-9 (2005).
31. Ikram, M.A. *et al.* The Rotterdam Study: 2018 update on objectives, design and main results. *Eur J Epidemiol* **32**, 807-850 (2017).
32. Mackey, D.A. *et al.* Twins eye study in Tasmania (TEST): rationale and methodology to recruit and examine twins. *Twin Res Hum Genet* **12**, 441-54 (2009).
33. Spector, T.D. & Williams, F.M. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet* **9**, 899-906 (2006).
34. Klein, R., Lee, K.E., Gangnon, R.E. & Klein, B.E. The 25-year incidence of visual impairment in type 1 diabetes mellitus the wisconsin epidemiologic study of diabetic retinopathy. *Ophthalmology* **117**, 63-70 (2010).
35. Raitakari, O.T. *et al.* Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* **37**, 1220-6 (2008).
36. Hughes, K., Aw, T.C., Kuperan, P. & Choo, M. Central obesity, insulin resistance, syndrome X, lipoprotein(a), and cardiovascular risk in Indians, Malays, and Chinese in Singapore. *J Epidemiol Community Health* **51**, 394-9 (1997).
37. Foong, A.W. *et al.* Rationale and methodology for a population-based study of eye diseases in Malay people: The Singapore Malay eye study (SiMES). *Ophthalmic Epidemiol* **14**, 25-35 (2007).
38. Lavanya, R. *et al.* Methodology of the Singapore Indian Chinese Cohort (SICC) eye study: quantifying ethnic variations in the epidemiology of eye diseases in Asians. *Ophthalmic Epidemiol* **16**, 325-36 (2009).
39. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. FUMA: Functional mapping and annotation of genetic associations. *bioRxiv* (2017).
40. Consortium, G.T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
41. Zhernakova, D.V. *et al.* Identification of context-dependent expression quantitative trait loci in whole blood. *Nat Genet* **49**, 139-145 (2017).
42. Booij, J.C. *et al.* Functional annotation of the human retinal pigment epithelium transcriptome. *BMC Genomics* **10**, 164 (2009).
43. Li, M. *et al.* Comprehensive analysis of gene expression in human retina and supporting tissues. *Hum Mol Genet* **23**, 4001-14 (2014).
44. Young, T.L. *et al.* Whole genome expression profiling of normal human fetal and adult ocular tissues. *Exp Eye Res* **116**, 265-78 (2013).

45. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
46. Frey, B.J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972-6 (2007).
47. Pers, T.H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).
48. Bodenhofer, U., Kothmeier, A. & Hochreiter, S. APCluster: an R package for affinity propagation clustering. *Bioinformatics* **27**, 2463-4 (2011).
49. Saito, R. *et al.* A travel guide to Cytoscape plugins. *Nat Methods* **9**, 1069-76 (2012).
50. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504 (2003).
51. Beaulieu, J.M. & Gainetdinov, R.R. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacol Rev* **63**, 182-217 (2011).
52. Gardner, M., Bertranpetit, J. & Comas, D. Worldwide genetic variation in dopamine and serotonin pathway genes: implications for association studies. *Am J Med Genet B Neuropsychiatr Genet* **147B**, 1070-5 (2008).
53. D'Souza, U.M. & Craig, I.W. Functional polymorphisms in dopamine and serotonin pathway genes. *Hum Mutat* **27**, 1-13 (2006).
54. Verhoeven, V.J. *et al.* Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat Genet* (2013).
55. Kiefer, A.K. *et al.* Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. *PLoS Genet* **9**, e1003299 (2013).
56. Pickrell, J.K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-17 (2016).
57. Brown, B.C., Asian Genetic Epidemiology Network Type 2 Diabetes, C., Ye, C.J., Price, A.L. & Zaitlen, N. Transethnic Genetic-Correlation Estimates from Summary Statistics. *Am J Hum Genet* **99**, 76-88 (2016).